"In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institution shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

THE APPLICATION OF DEMAND THEORY TO THE DECISION MAKING PROCESS IN A SERVICE INDUSTRY

A THESIS

Presented to

The Faculty of the Graduate Division

by

Francis Leonard Wiegmann

In Partial Fulfillment

of the Requirements for the Degree of

Master of Science in Industrial Engineering

Georgia Institute of Technology

June, 1958

THE APPLICATION OF DEMAND THEORY TO THE DECISION MAKING PROCESS IN A SERVICE INDUSTRY

5R
12T

Approved By:

Date of Approval: May 13, 1958

ACKNOWLEDGEMENTS

The writer expresses his appreciation to those who have helped make this study possible, especially to Dr. David C. Ekey for his interest, direction, and assistance in completing this thesis. For the wholehearted cooperation of the management of the Atlanta Transit System, Incorporated, without which this work could not have been done, the author is extremely grateful. A word of appreciation is also due Dr. Joseph Krol and Dr. James W. Walker for their advice and participation in this research. In addition, the writer is indebted to the Rich Electronic Computer Center for the use of the IBM 650, and to the Graduate Division of the Georgia Institute of Technology for the research fellowship under which this research was conducted.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

As businesses and their associated financial risks grow larger management strives to improve its decision making process. The industrial engineer assists management in the evaluation of the effects of its decisions. It is, therefore, advantageous for him to have available a decision tool which can quantitatively evaluate the effects of changes in operating policy on the demand for the company's products.

This study investigates the methods of approach of econometricians, who have been concerned with demand problems in the past. The work of Professors Moore and Schultz provides a foundation for the investigation of demand for intra-city transit services in Atlanta.

Observations of nine variables are presented and adjusted appropriately for population growth and for change in the purchasing power of money. After investigation of the proper form of the demand function with a graphic scaffolding technique, an initial multiple regression model is developed to determine the significance of the time variables. Multiple regression models are then developed for the 1936 to 1952 period (omitting the war years) with additional variables to reduce the effects of the time factors. These models are evaluated by two methods in terms of their ability to predict demand and to simulate the evaluation of changes in operating policy for the 1953-1957 period. Additional multiple regression models are developed by similar procedures to provide

up-to-date decision tools for the 1936-1957 period (omitting the years 1942-1946). With the use of the models price elasticities of demand, service elasticities of demand, rates of shift of the demand curves, and the significance of the variables are studied.

The econometricians approach to demand theory is shown to be of value to the industrial engineer when studying demand problems on the company level. The relative value of multiple regression models as decision tools is presented. Conclusions based on the results are made concerning the transit market in Atlanta, and recommendations are made for further research.

CHAPTER I

INTRODUCTION

DEMAND ANALYSIS--A DECISION TOOL

As a business in this country and the associated financial risks grow larger management constantly searches for greater quantities of accurate operating statistics. This information is required to be as current as possible to predict future activities and to provide improved confidence limits in planning. It is the responsibility of the industrial engineer to assist management in the evaluation of the probable effects of changes in operating policies on the demand for their products. It is desirable for the industrial engineer to have available a decision tool, or tools, which could quantitatively evaluate the effects of these changes in operating policy.

In the past engineers have used various methods to solve these forecasting problems. Some have relied upon their sense of judgment, their experience, and their intuition to estimate these effects. More sophisticated approaches have included methods of extrapolation of demand curves. However, if the interrelationships among the significant factors in the demand functions can be understood, management can anticipate changes in demand brought about by its decisions.

The economist and econometrician have, in the past, been concerned with problems of supply and demand. Most of the studies resulting have been concerned with national markets and not with the demand problems of an individual enterprise. It is one aim of this study to investigate the methods of the econometrician and to determine whether such methods are applicable to demand problems on the company level.

The industrial engineer, through the discipline of operations research, has in recent years been placing greater emphasis on the development of mathematical models. Although the type of model investigated in this study has not been used as frequently as it might be, it will probably be used more when its advantages are more fully understood by the industrial engineer. This investigation indicates the relative value of multiple regression models as decision tools, when time is an important factor.

Knowledge of the foregoing ideas led to the following hypothesis: The theory and measurement of demand as set forth in econometrics is applicable to company level decision making and is of value as a method of approach, in industrial engineering and operations research, to the solution of demand problems that confront management.

The classical study of demand by econometricians is generally restricted to tangible products in national markets such as those for agricultural commodities, steel, and automobiles. In addition to demand problems at the company level, the writer has elected to study a non-tangible service type industry. The objective is to determine the applicability of demand theory under these restrictions.

The subject chosen for this research involves demand functions for intracity transit services, i.e., buses and trackless trolleys. The influence of certain factors on demand for this service is evaluated. Multiple regression models, recommended by some econometricians, are used to develop decision models. These models are evaluated in terms of their use as prediction tools.

Literature Survey

In the attempt to draw on the disciplines of other fields to solve demand problems in industrial engineering, considerable study was required in econometrics and mathematical statistics. The literature survey encompassed all issues of Econometrica, a journal which represents a prime source of information concerning demand functions, supply functions, production functions, and other related topics.

Several important facts were found from this survey. There are two basic sources of data for the construction of demand functions: market data (time series) and family-budget data. Professor Henry L. Moore, Professor Henry Schultz, Professor Wassily Leontief, and Professor A. C. Pigou are prominent authors associated with the time series approach. Professor Ragnor Frisch, Dr. Jacob Marschak, Professor

A. C. Pigou, and Professor René Roy are prominent authors advocating the approach through family-budget data. There are three statistical procedures developed by Professor Moore: the method of link relatives, the method of trend ratios, and the method of multiple correlation. The most recent work on demand theory, which deals with many aspects of the subject, is by Henry Schultz. Schultz accepts Moore's theories and proceeds to quantitatively evaluate numerous alternative procedures in this field. Professor Schultz's book, The Theory and Measurement of Demand, is a very complete reference on demand theory and was used extensively during the course of this research. The statistical procedures used in this thesis are those treated by Schultz, but other statistical references are employed. Since computations were done on the I.B.M. 650 electronic computer, reference was also made to the multiple regression subroutines and their associated techniques.

This background material is discussed further, where appropriate, throughout this thesis.

Historical Development

The law of demand, or the law of price which embraces it, was commonly stated a century ago as: "The price of goods varies directly as the quantity demanded and inversely as the quantity supplied."[1] In symbols,

---

[1]Baptiste, Jean, Catéchisme d'economic politique, 4th Edition, Paris, 1835, chapter xi, p. 104.

$$P \propto \frac{D}{S} \qquad (1)$$

where P is the Price; D is the quantity demanded; and S is the quantity supplied. This is not a very precise statement, since quantity demanded actually means quantity demanded at a certain price, and quantity supplied means quantity supplied at a certain price. A customer does not usually "demand" until he knows the price; and the supplier does not "supply" until the price is set. In 1838, Cournot stated the law of demand in another form as follows: " . . . the sales or demand D is, for each article, a particular function F(p) of the price p of such article. To know the form of this function would be to know what we call the law of demand or of sales."[2] Mathematically Cournot's law is

$$D = F(p) \qquad (2)$$

Alfred Marshall developed this concept more fully and popularized it. It has been called the Cournot-Marshall law of demand.[3] Marshall states his general law as: "The greater the amount to be sold, the smaller must be the price at which it is offered in order that it may find purchasers; or, in other words, the amount demanded increases with a fall in

---

[2]Cournot, Augustin, Researches into the Mathematical Principles of the Theory of Wealth, translation, Bacon, New York, 1897, p. 47.

[3]Marshall, Alfred, Principles of Economics, 8th Edition, London, 1920, Book III.

price, and diminishes with a rise in price." However, there is nothing "general" nor "universal" about this law. It is true under certain circumstances, and the opposite is true under others.

It is not always an easy matter to find the form of the function F(p). One could conduct an experiment in which customers are asked how much they would buy at a given price. All the customers would not be able to answer the questions without knowing the prices of other products as well as the price of the product in question. Leon Walras in 1873 stated these ideas mathematically as follows:

$$D = F(p_1, p_2, p_3, \ldots, p_n) \tag{3}$$

where D is the quantity demanded, p_1 its price, and p_2, p_3, . . . , p_n the prices of all other commodities.[4] This is the law of demand of the Lausanne School. If we first take all the variables into consideration, then assign constant values to all variables except the price and quantity of the commodity in question, we obtain the Cournot-Marshall law as a special case from the Walras law. Note here that the demand curve will depend on the constants assigned to the subsidiary variables (i.e., the prices of all other commodities). The idea of introducing the prices of all other products into

---

[4]Walras, Leon, Elements d'economie politique pure, 4th Edition; Lausanne and Paris, 1900, pp. V-vii and especially M. 1, p. vii; also, pp. 122-33 and 208-15.

the equation and then assigning constant values to them is a marked improvement over the classical and neo-classical conception of the demand curve.

The way, then, to deduce the demand function is to observe the behavior of consumers in mass markets. Each market transaction will represent a unique set of conditions, and will give us a single point on our demand "curve" or surface. Therefore, to obtain the form of the function we must have numerous observations covering a considerable period of time. As a result, the variables dealt with are functions of time; and the law becomes

$$D = F(p_1, p_2, p_3, \ldots, p_n, t) \qquad (4)$$

with the addition of the time variable added to equation (3).

Another approach to the problem in terms of Indifference Curves and Indices of Utility with related Contract Curves was introduced by F. V. Edgeworth in 1881.[5] He was the first to write the utility of a commodity as a function not only of its quantity but also of other quantities. V. Pareto and E. Slutsky (1915) made further contributions to this aspect of demand theory. From their efforts came the general market demand function:

$$X = F(y_1, y_2, y_3, \ldots, y_n, R) \qquad (5)$$

---

[5] Edgeworth, F. Y., Mathematical Psychics, London, 1881, p. 31.

where X is the quantity demanded, y_1 its price, $y_2, \ldots,$ y_n the prices of other commodities; and R represents the size and distribution of income in the economy. Schultz concludes,

> With the development of his general law of demand Pareto has corrected, completed, and extended the work of Walras and others on the relation of utility to demand; with the explicit introduction of income into the demand function Slutsky and, later, Hicks and Allen have rendered a similar service to Pareto.[6]

The requirement of numerous observations, usually covering a considerable period of time which involves dynamic changes in the market, leads to the restatement of equation (5) as follows:

$$X = F(y_1, y_2, \ldots, y_n, R, t) \qquad (6)$$

where X is the quantity of the commodity demanded, y_1 its price, $y_2, y_3, \ldots, y_n$ the other prices or influencing factors, R the size and distribution of income, and t is time -- a "catch-all" for the resultant of those factors which cannot conveniently be measured separately, but which change slowly and smoothly.

Study of the dynamics of demand requires a knowledge of the direction of change of the system (given by statistical equations) and also its velocity. Thus, the quantities consumed, as well as other variables, must be considered as

---

[6]Schultz, Henry, <u>The Theory and Measurement of Demand</u>, Chicago, 1938, p. 50.

vectors which are functions of time.

Sir Issac Newton developed laws of motion and conservation of momentum which are useful in studying the dynamic problems of physical systems. Unfortunately, there are no economic laws which parallel the work of Newton in studying dynamic problems of demand. We encounter even greater difficulty in predicting how a demand curve will move in the future. The best we can do at the present state of our knowledge is to make a study of statistical equilibria, isolate their routine of change, if it exists, and hope that this routine will continue to operate in the future. Such a routine of change is represented by the dynamic demand function

$$X = F(y_1, y_2, y_3, \ldots, y_n, R, t) \qquad (7)$$

By its nature it cannot have the same heuristic properties as the laws of mechanics.

Statistical Procedures in Use

It is impossible to derive a demand curve from statistics without making some assumptions regarding the nature of the theoretical function and the interrelations of the variables. One general hypothesis is made, that the unknown demand curve can be approximated by empirical equations. A more specific hypothesis is that the available data falls into one of two main classes --time series of prices and quantities, and family-budget data. In time series, the dynamic demand function is used, and one assumes that tastes remain constant or that they change regularly and smoothly with time. One then makes assumptions as to the variables which should be included in the function. In family-budget data, one is faced with the problem that such data are not generally available for consecutive periods. One, then, is only able to find the demand function at a single point in time. It is impossible to use the dynamic demand function with family-budget data, and one must use the static demand function.

$$X = F(y_1, y_2, \ldots, y_n, R) \tag{8}$$

where the variable time is not present. Other assumptions are made regarding the comparability of families; and, as before, about the form and specific variables of the function.

Market Data (Time Series) Methods

Professor Moore's methods.--Schultz attributed the statistical study of demand solely to Professor Henry L. Moore, even though Marshall (1885), Cournot, Pigou (1910), Tschayanow (1912) and others preceded him in the development of some aspects of the problem. In 1914 Moore published his *Economic Cycles: Their Law and Cause*.[7] Moore's procedures have been used by Dr. Mordecai Ezekiel, Mr. L. H. Bean, Professors G. H. Warren, F. A. Pearson, Holbrook Working, and

---

[7]New York, 1914.

Dr. E. J. Working. Moore's work has also influenced price analysis done by statisticians for federal and state governments. Professor Moore's contributions to the solution of the demand problem are three:

1. He restated the hypothetical, statical law of demand in a form admitting of concrete, inductive treatment.
2. He devised ingenious statistical techniques, such as the method of link relatives and the method of trend ratios, for handling the time variable, and was among the first to apply the method of multiple correlation to the study of demand.
3. He succeeded in deducing for the first time the statistical demand curves for several important commodities, and in measuring their elasticities of demand.[8]

Professor Moore tacitly assumed that there exists a routine in the demand behavior of human beings in the market, that statistical data of consumption and prices will reflect this routine, and that unknown theoretical demand functions can be approximated by various fitted empirical curves. The first assumption implies that there have been no significant changes in tastes and desires in the consumers, so that the dynamic demand function will hold for the entire period. An abrupt change in the middle of the period under study might make it difficult to derive an accurate function.

---

[8] Schultz, <u>op. cit.</u>, p. 65.

The second assumption implies the existence of more than one equilibrium position for supply and demand within the period. The third assumption requires that the demand curve and its derivatives shall be continuous within the range of interest.

Moore treated the demand problem statistically by several methods: the method of multiple correlation, the method of relative changes (link relatives), the method of trend ratios, and combinations of these.

The method of relative changes consists of finding the functional relationships not between the absolute prices and absolute quantities but between the relative changes in the price of the commodity and the relative change in the quantity demanded. As a measure of the relative change we may take either the percentage change in the value from one year to the next or the ratio of the given year's value to that of the preceding year (link relatives).

The method of trend ratios derives the demand curve from the ratios of prices and quantities to their respective trends, and not from the absolute prices and corresponding absolute quantities. By the use of ratios an attempt is made to eliminate the disturbing effects of secular trends.

After Moore adjusted his data he tried to approximate the law of demand, first by linear functions:

$$X = F(y_1, \ldots, y_n) \qquad (9)$$

$$X = a_o + a_{11}Y_1 + a_{12}Y_2 + \ldots + a_{1n}Y_n$$

Thus, multiple correlation was used to determine a_o, $a_{11}, \ldots, a_{1n}$. In equations (9) time t is not present, since each factor has been adjusted for trends in some way. B. B. Smith showed that the introduction of time also explicitly in (9) would greatly improve the accuracy of the estimate of the dependent variable.[9]

Professor Leontief's method.--The chief feature which differentiates Leontief's method from all others is his attempt to derive elasticities of both demand and supply with one calculation from the same set of unadjusted statistics of prices and corresponding quantities. He claims that it is unnecessary and undesirable to make allowances for changes in purchasing power of money, prices of substitutes, and other disturbing factors. There is no place in his methods for adjusted data, link relatives, trend ratios, the "lag" method, multiple or partial correlation and other devices used by Moore.

Leontief's underlying assumptions are:

1. Each transaction represents the intersection of an instantaneous demand curve with a theoretical

---

[9]Smith, B. B., "The Error in Eliminating Secular Trends and Seasonal Variation Before Correlating Time Series," Journal of American Statistical Association, XX (1925), pp. 543-545.

instantaneous supply curve, which change positions from time to time.

2. For each of these curves the elasticity is approximately constant.
3. The shiftings of the demand and supply curves are independent of one another, and do not affect the shape (elasticity) of the curves.

The first assumption implies that we must not only determine elasticities but also the extent to which the curves have shifted. The second assumption implies that the demand and supply curves appear as straight lines on double logarithmic paper.

Leontief's method consists of fitting two straight lines to the scatter of logarithms of the observations. If one of the curves be given, the other can be found from it. By taking as the demand the line that which minimizes the sum of the squares of the deviations about it measured parallel to the supply curve, and vice versa, Leontief derives a pair of curves. But neither curve is given, and in order to get a unique set of curves, the observations are arranged in chronological order, divided into two parts, and each half is fitted with a pair of curves. Then the criterion is imposed that the two scatters should have one pair of curves in common. These two curves represent average demand and supply curves.

By making the conventional assumption that the

Cournot-Marshall demand curves shift in any direction independently, Leontief throws away the fundamental principle of the theory of equilibrium, i.e., that the demand for any one commodity is a function of its price, and all other prices.[10] His statistical procedures are limited in many ways, enumerated by Schultz.[11]

Professor Pigou's second method.--Professor Pigou published his first method dealing with family-budget data in 1910.[12] It will be considered later.[13] The second method was published in 1930.[14]

Pigou is critical of both Moore's and Leontief's methods. He does not, however, refer to Moore's multiple correlation. Like Leontief, Pigou works with the Cournot-Marshall demand curve.[15] He defines demand curves as "the quantities of a commodity that a market will buy during a

---

[10]Leontief, Wassily, "Ein Versuch zur statistichen Analyse von Angebot und Nachfraze," Weltwirtschaftliches Archiv, XXX, Heft 1 (July, 1929), pp. 1-53.

[11]Schultz, op. cit., pp. 93-95.

[12]Pigou, A. C., "A Method of Determining the Numerical Value of Elasticities of Demand," Economic Journal, XX (1910) pp. 636-40; and reprinted in his Economics of Welfare (London 1920).

[13]Infra, p. 18.

[14]Pigou, A. C., "The Statistical Derivation of Demand Curves," op. cit., pp. 384-400, and reprinted in Pigou and Robertson, Economic Essays and Addresses, London, 1931, pp. 62-83.

[15]Ibid., p. 384.

short interval, say a year, in response to different average prices proper to the interval." He also points out that the quantity demanded is a function not only of its price but also of "the conditions of supply of several other things," but he assumes these things to be constant in each interval.

Pigou's assumptions are:

1. The demand curve for each interval is a curve of constant elasticity.
2. The rate of shift is such that the distance between the first and the second position (on a logarithmic scale) is the same as between the second and third.

Pigou's statistical procedure makes use of the geometric proposition that through three noncollinear points only three triads of parallel straight lines can be drawn which are equidistant from one another vertically (which also implies horizontally). If, however, we specify which point shall lie on the middle line, then only one triad is possible. This procedure is limited in many ways and is not as desirable as others.[16]

Professor Schultz's method.--Professor Schultz follows Professor Moore basically but has improved on his methods. He presents and evaluates them in the most scholarly fashion. Schultz's assumptions can be summarized as follows:

1. There exists a routine in the demand behavior of human beings.

---

[16] Schultz, *op. cit.*, pp. 102-104.

2. The statistical data of consumption and prices are such as to reflect this routine of demand.
3. The unknown theoretical demand function can be approximated by various empirical curves. Other assumptions concerning the adjustment of data, the form of the function, and the method of curve fitting are made.

Schultz was careful to evaluate the difference between the trend ratios, link relatives, and multiple regression (time series) methods, as well as the advantages of adjusted data by actually computing the same demand functions by each method. Before his death (1937), Schultz derived the demand functions and the elasticities of demand for corn, cotton, hay, wheat, sugar, potatoes, oats, barley, rye, buckwheat, pork, beef, veal, mutton, and lamb. He also derived the demand functions for Canadian sugar, tea, and coffee. His methods usually start by adjusting the data for changes in population and the purchasing power of money. The period for study is then broken into homogeneous sub-groups with regard to tastes. Each sub-period is analyzed separately by first using a "graphic scaffolding" to find the appropriate form of the function. Multiple regression is then used to derive the dynamic demand function. The methods described here are judged by the writer to be the most reasonable and useful of all the methods studied.

Family-Budget Data Method

Professor Pigou's first method.--This method is derived from the theory of utility and makes use of budget data. Pigou assumes that the degree of utility of money is independent of the degree of utility of any commodity. He also assumes that the tastes and temperament of the people in any two adjacent groups are approximately the same.

Even if the assumptions are allowed, this procedure can yield only the ratio of two elasticities of demand. To obtain absolute elasticities, another method must be used. However, the assumption is questionable that

> Since a small change in the consumption of any ordinary commodity on which a small portion of a man's total income is spent cannot involve any appreciable change in the marginal degree of utility of money to him, the elasticity of the utility curve . . . is equal to the elasticity of the demand curve . . . .[17]

Professor Frisch's method.--Professor Frisch's methods have as their object the measurement of the degree of utility of money and not the derivation of statistical demand curves. But the method yields, as a by-product the Cournot-Marshall demand curve. Frisch's main work was published in 1932.[18] The basis for his procedures is the formula

$$\frac{u_1(a)}{p_a} = \frac{u_2(b)}{p_b} = \frac{u_3(c)}{p_c} = \ldots \qquad (10)$$

---

[17] Schultz, op. cit., p. 110.

[18] Frisch, Ragnor, New Methods of Measuring Marginal Utility, ("Beitrage zur okonomischev Theorie," No. 3, Tubingen, 1932).

where $u_1(a)$, $u_2(b)$, . . . are the degree of utility of quantities a, b, . . . of commodities A, B, . . . ; and P_a, P_b, . . . are the prices. Equation (10) states that at equilibrium the individual distributes his expenditures in such a way that a dollar's worth of any commodity yields the same utility as a dollar's worth of any other commodity. Frisch then proceeds to derive an equation which expresses the relation between the degree of utility of money and the degree of utility of the commodity of comparison, which he calls the surface of consumption. When data are available for the construction of a surface of consumption, and when the demand for the commodity in question is practically independent of the demands for other goods, there is no doubt that this procedure will yield the Cournot-Marshall demand curve for the particular income level. This method, however, is not considered desirable because of its use of family-budget data.

Dr. Marschak's method.--Dr. Marschak's method[19] is summarized as follows.[20] Family-budget data give the relation between money incomes of various households during a specified period of time, the quantities of the various goods and services, and the prices paid for these goods and services. An empirical formula connects these categories for households of

---

[19] Marschak, Jacob, Elastizitat der Nachfrage ("Beitrage zur okonomischen Theorie," No. 2, Tubungen, 1931.

[20] Schultz, op. cit., pp. 117-18.

a fixed size. The quantity consumed may then be varied by varying either the price, the income, or both. Then a single market demand curve is derived, for which it is assumed that all prices rise and fall together. The final market demand curve is the sum of the demand curves of the individual households. In brief, Marschak's demand curve is first an income curve. What he measures is the elasticity of quantity with respect to income, not with respect to price. Although this method has some advantages as well as disadvantages, it still derives demand curves only for the case when prices and quantities rise and fall together.

Professor Roy's method.--Professor Roy's fundamental contribution[21] is the investigation of the relation which should exist between the distribution of income and the law of demand for the group of first necessities, for the group of commodities other than first necessities, and for individual commodities. He appeals to the relation existing between Pareto's law of distribution of income and the law of demand for all commodities taken as a group. Roy assumes that during the period covered there have been no changes in population size, or in size and distribution of income. He further assumes that each individual has a scale of preferences for various commodities. If income diminishes

---

[21]Roy, René, "La Demande dans ses rapports avec la répartition des revenues," Metron, VIII, No. 3 (1930), pp. 101-153; "Les Lois de la demande," Revue d'économie politique, XLV (1931), pp. 1190-1218.

while prices remain constant, an individual will forego the use of those commodities which he considers least indispensable. Professor Roy's approach, and Dr. Marschak's is essentially the one suggested by Pareto in 1895,[22] although apparently neither man was aware of Pareto's contribution. Since Pareto's law does not strictly hold for low income groups, and since adequate data are not available to evaluate Roy's method properly, it may be less desirable than some other approaches. However, it does have merit; and when better data are available, it may be useful in calculating the Cournot-Marshall demand curve.

The foregoing briefly describes the various approaches which have been suggested in the solution of demand problems. It is recognized that one must make assumptions about the demand curve itself before deriving it. If one wants to know the probability that a given event had its origin in a given cause, it is necessary to first determine the probability that the cause exists. Since this probability is not generally known, one must make plausible assumptions concerning it.

The time-series approach suffers from difficulty in disentangling the "time factor" from the data. It is not suitable for deriving the elasticities of demand for classes of commodities and services, such as clothing, housing, and

---

[22] Pareto, Vilfredo, "La Legga della domando," Giornale degli economisti, X (2nd series, 1895), pp. 59-68.

amusements. It is applicable only when certain supply conditions prevail, but has the advantage of enabling one to study shifting demand curves.

The budget-data approach also has severe limitations. It does not permit derivation of elasticities for intermediate goods, or goods consumed directly by human beings. It does not yield a measure of the rate of shift of the demand curve and does not adapt itself to analysis of the effects of substitutes. The budget-data approach, however, has the advantage that its validity is not dependent on the mode of shift of the supply curve.

It is apparent that there has been considerable scholarly study in the demand theory phase of econometrics. These objective, quantitative efforts should inspire the industrial engineer to similar achievements with demand functions at the company level of operation.

CHAPTER II

PROCEDURES

Underlying Theory

The basic demand function and its underlying theory, which is hypothesized as applicable to the dynamics of transit service demand, is as follows:

$$X = F(y_1, y_2, \ldots, y_n, R, t) \tag{11}$$

where X is the demand for transit service; y_1 is the adjusted average fare; $y_2, \ldots, y_n$ are the effects of substitute means of transportation or other controlling factors; R is the adjusted effective buying income per capita for Atlanta; and t is time --reflecting those unknown factors which change slowly with time. This dynamic demand function requires three basic assumptions:[23]

1. There exists a routine in the demand behavior of transit riders.
2. The statistical data of revenue passengers, average fares, and other significant factors[24] are such as to reflect this routine of demand.
3. The unknown theoretical demand functions can be approximated by various empirical curves.

---

[23] Supra, p. 16.

[24] Infra, pp. 68-76.

The first assumption implies that a mass market is being studied. This assumption, used by the econometrician in national markets, is weakened in this thesis by applying it to the "transit market" in Atlanta.

Preliminary studies of this system disclosed that data for years before 1936 were not available, and influenced the selection of that year as the beginning of the period for this investigation. The selection of the time interval was further influenced by the cessation of manufacture of automobiles during the years 1942-1946. Since this cessation caused transit service demand to be unrealistically inflated, the World War II years (1942-1946) were not included in the following development.

The initial data for demand and price were collected partially from transit company records and partially from a report by the Georgia State Senate Transit Study Committee on a plan for improving transit service.[25]

Since the use of unadjusted data to develop a demand function will surely lead to a function which is affected by changes in population growth and purchasing power (e.g., inflationary trends) of money, it would be necessary, then, to introduce these two factors explicitly into the demand function as additional variables. However, if the data are

---

[25]A Plan For Transit Improvement in the Metropolitan Area of Atlanta, Georgia State Senate Transit Study Committee, and Simpson and Curtin, Transportation Engineers and Consultants, November 1953.

adjusted, these variables will not need to be present explicitly in the demand function. Reducing the number of variables, as much as possible, is advisable in the light of difficulties which arise in the statistical procedures when the number of variables is large. Therefore, all data for this study were adjusted appropriately.

Three adjustment series were used. The U. S. Department of Labor's Consumer Price Index for Atlanta, Georgia, Series A-2 (1947-1949 base) was used to adjust all money series. The Metropolitan Area Population Series, as published by the Atlanta Chamber of Commerce, was used to adjust the Atlanta Effective Buying Income Series. The Atlanta Population Served by Transit Series, as estimated by The Atlanta Transit System, Incorporated, was used to adjust the Demand Series and the Transit Vehicle Miles per Capita Served Series.

The two series, per capita demand, and adjusted average fare, are plotted in Fig. 1 on a logarithmic scale to investigate whether there have been any significant or abrupt changes in the customer's desire for transit service. It appears from the figure that significant changes in taste have taken place. It seems appropriate then, to analyze the 22 years from 1936 to 1957 in two separate periods. However, the number of years available after omission of the war years is so small that it was decided to proceed with one period instead of two. This point is further amplified in following

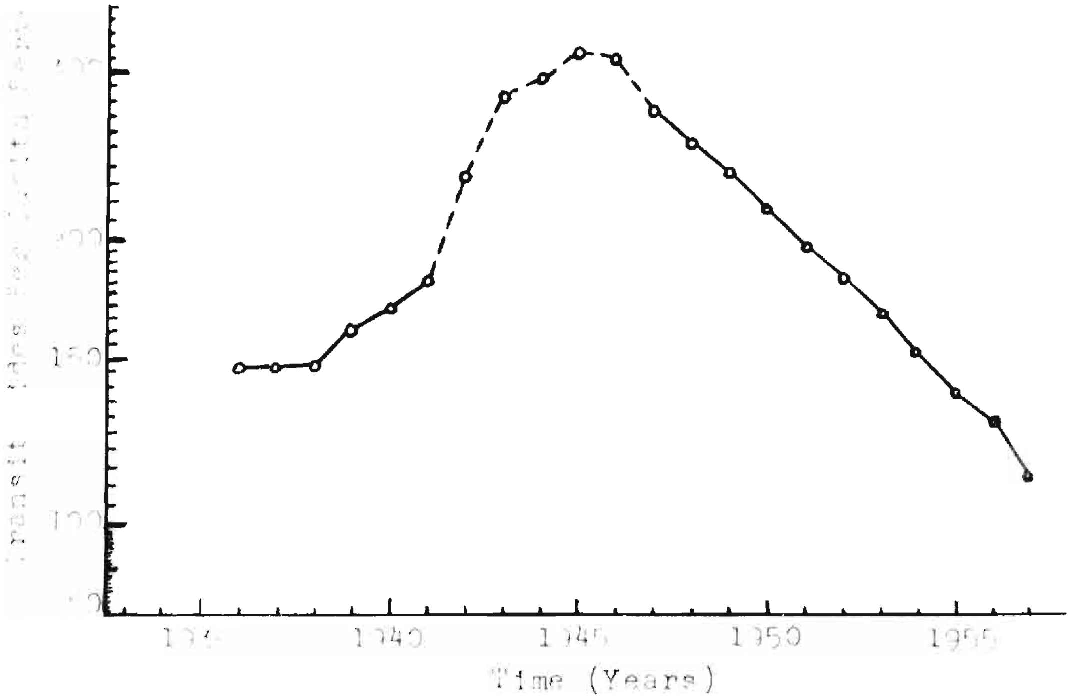


Fig. 1a: Basic Demand Series

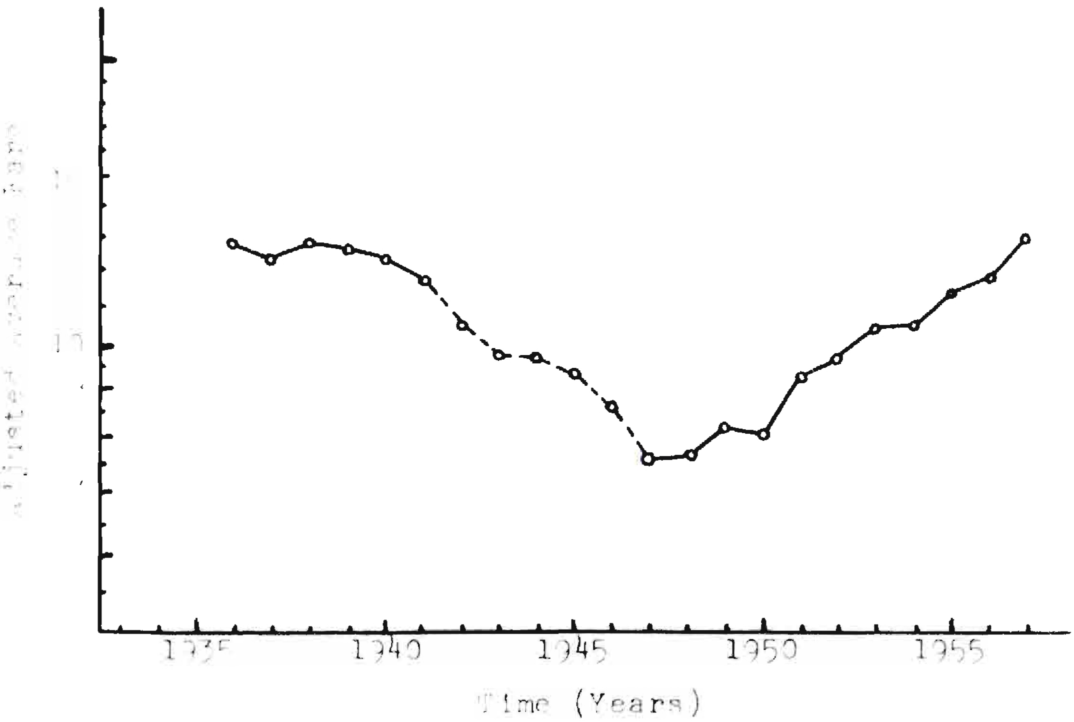


Fig. 1b: Basic Price Series

sections concerning "degrees of freedom" associated with the statistical procedures.[26]

Development of Models

To help us decide what the form of the demand function should be, Schultz suggests a "graphic scaffolding",[27] which will show how the demand curve is affected by long term secular trends. By plotting the logarithms of per capita demand on the ordinate, and the logarithms of adjusted average fare on the abscissa as in Fig. 2, a first approximation (AA) to the demand line can be fitted by eye. This demand line is drawn through the point whose coordinates are the means of the two series, with a slope which is representative of the slopes of the individual connecting lines. The deviation (PQ) of each plotted point from the demand line (AA) is then plotted from a zero reference line in order of time in Fig. 3. The trend curve BB is then fitted by eye to this plot of deviations against time. If this curve has zero slope, there will be no significant secular trends. If the curve is linear and at some slope other than zero, the variable t (time) will be represented in the demand function. If the curve is parabolic, both the variables t^2 (time squared) and t (time) will be included in the demand function. As shown in Fig. 3, this will be the case for

---

[26]Infra, pp. 37-38.

[27]Schultz, *op. cit.*, pp. 184-186.

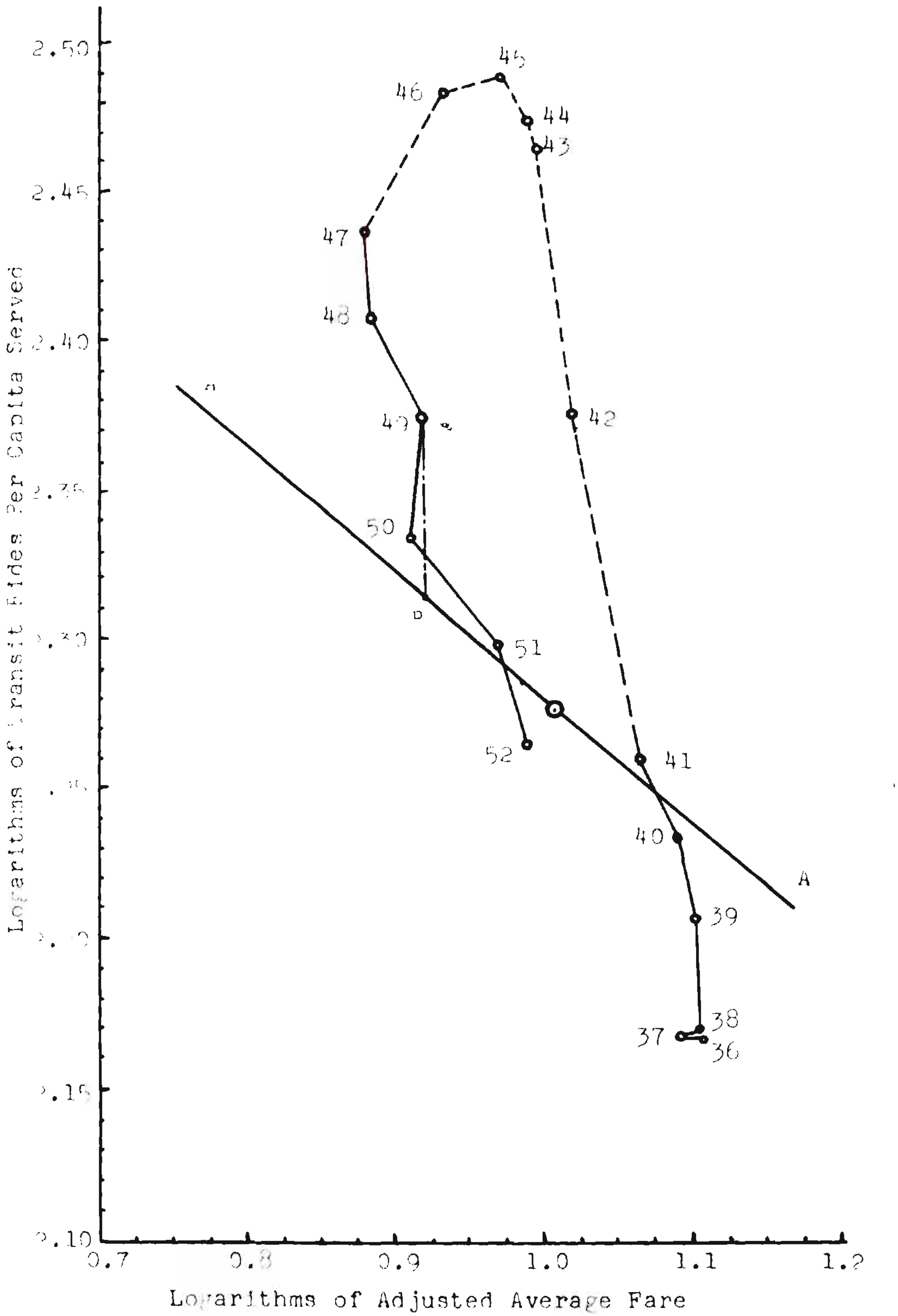


Fig. 2: Demand Not Adjusted For The Effects of Time

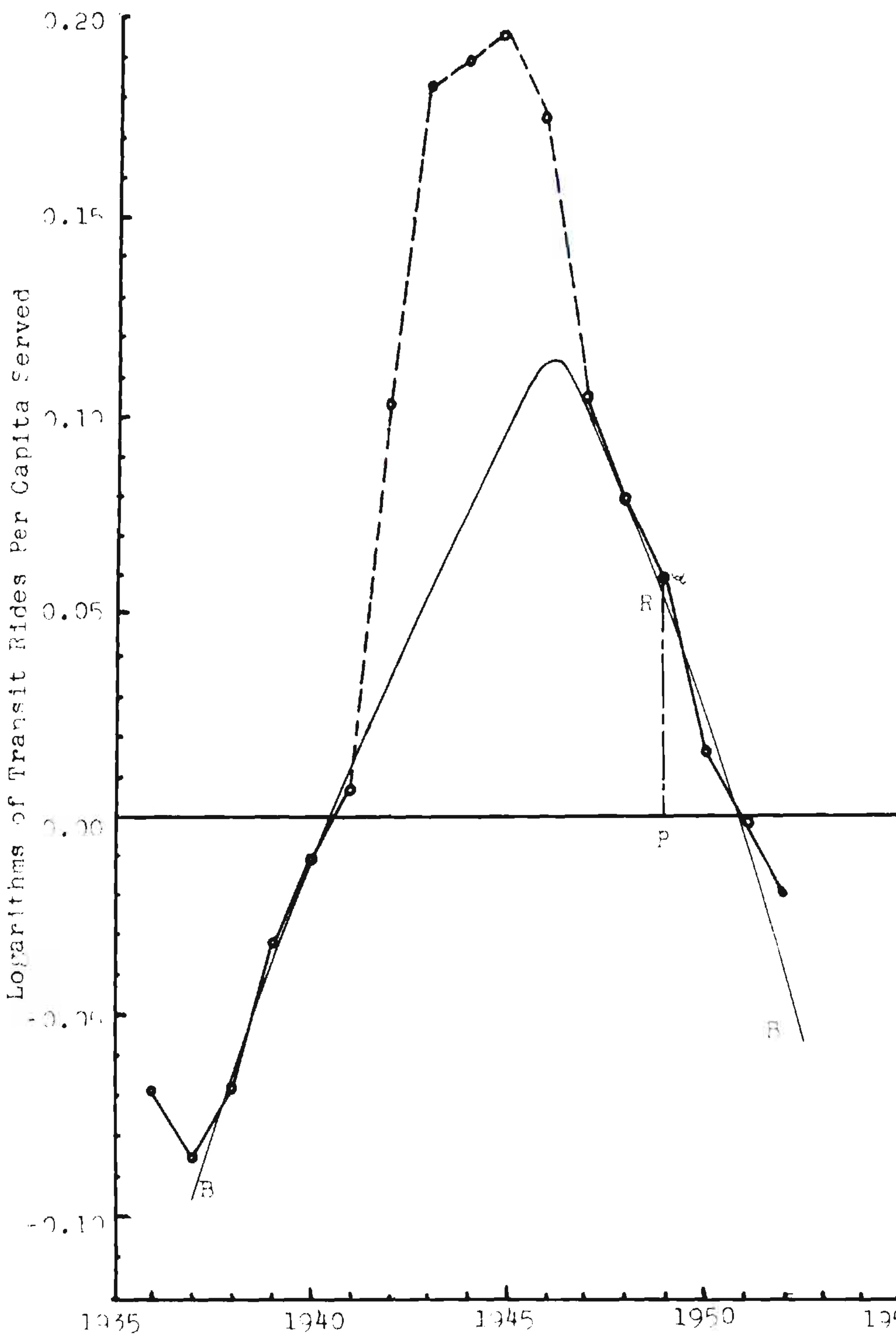


Fig. 3: Relation Between Demand and Time

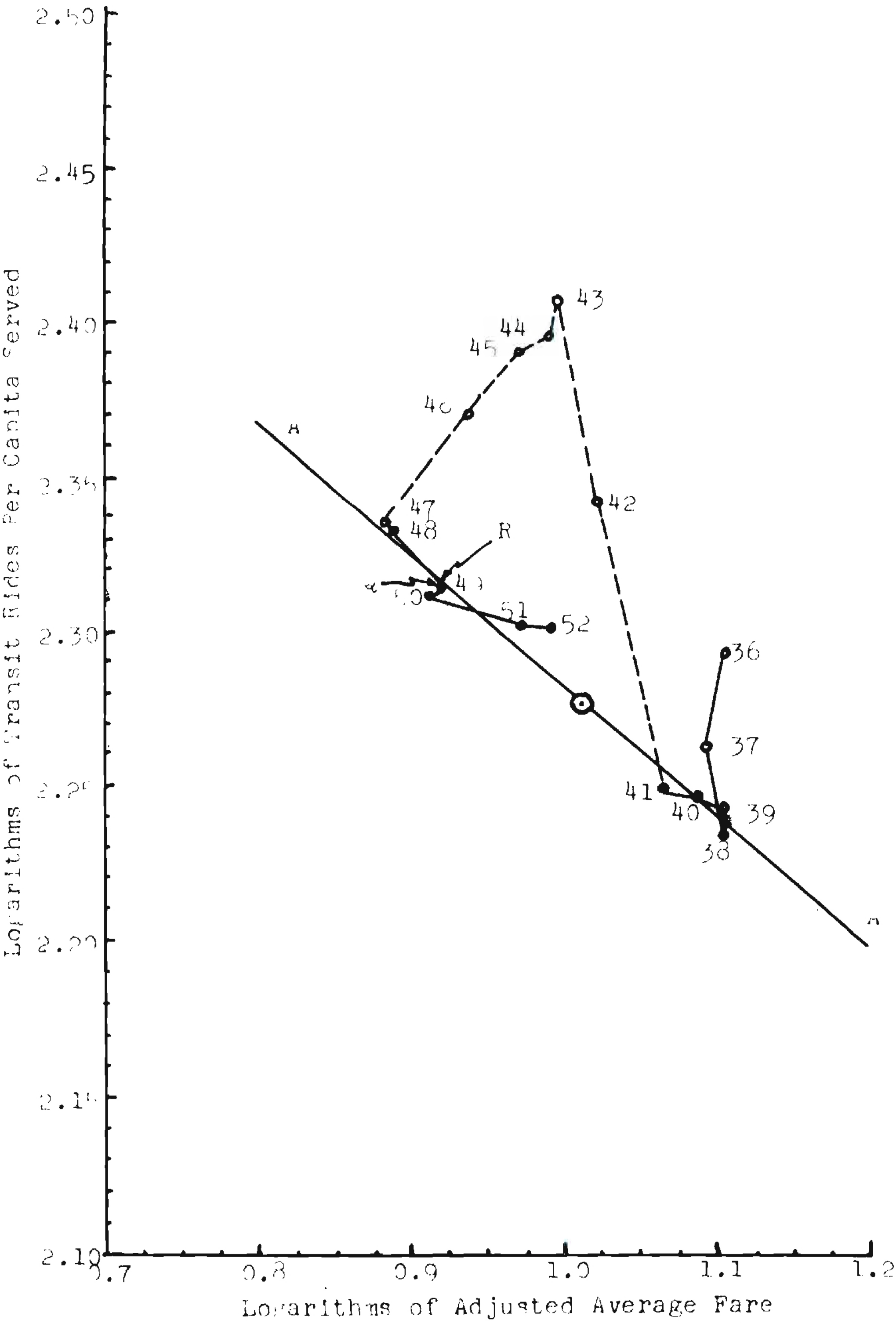


Fig. 4: Demand Adjusted For The Effects of Time

Model I. Since logarithms of the data were used, this exploratory model will be linear in its logarithmic form as in equation (12) or multiplicative when transformed as in equation (13).

$$\log X = \log b_0 + b_1 \log y + b_2 Mt + b_3 Mt^2 \qquad (12)$$

or

$$X = b_0 Y^{b_1} e^{b_2 t + b_3 t^2} \qquad (13)$$

By adjusting for the secular trends we can see that AA is a good first approximation to the true demand line in Fig. 4. For example, as is illustrated for the year 1949,

$$\overline{PQ} - \overline{PR} = \overline{RQ} \; . \qquad (14)$$

The same calculation is made for each plotted point resulting in a scatter which is concentrated about AA, the demand line. Thus the effects of secular trends have been eliminated from the consumption data.

The four series that will be used, for the years 1936-1952, to derive Model I are, therefore:

X_1 = log X = logs of per capita demand

X_2 = log y = logs of adjusted average fare

X_3 = t = time (1936 origin)

X_4 = t^2 = time squared (1936 origin)

M = $\log_{10} e$

and the form of the equation will be as in equation (12). The regression coefficients b_0, b_1, b_2, b_3 and their standard errors will be calculated by a modified Gauss' method. This method makes possible the simultaneous determination of all the elementary regression equations and their standard errors, and, of course, the regression coefficients and their standard errors.[28] The calculations are also simplified by normalizing each of the variables, i.e., subtracting the mean and dividing by the standard deviation of the series. After normalizing, the first step involves the calculation of the correlation coefficient between each combination of pairs of variables. Thus, for Model I, using equation (15):

$$r_{ij} = \frac{\dfrac{\sum X_i X_j}{n} - M_i M_j}{\left(\dfrac{\sum X_i^2}{n} - M_i^2\right)^{\frac{1}{2}} \left(\dfrac{\sum X_j^2}{n} - M_j^2\right)^{\frac{1}{2}}} \tag{15}$$

where X_i and X_j are the variables in question, M_i and M_j are their means, and n is the number of observations, we can construct the symmetrical intercorrelation matrix, shown in Fig. 5.

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{pmatrix}$$

Fig. 5: Symmetrical Intercorrelation Matrix

---

[28]Schultz, <u>op. cit.</u>, Appendix C.

The inverse of the intercorrelation matrix is determined by standard row transformation procedures. Most texts in statistics or matrix theory treat these methods. A good presentation may also be found in Schultz's Appendix C. The result of this inversion process is checked by matrix multiplication to obtain the unit matrix, i.e:

$$(r_{ij})\ (P_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{16}$$

where the matrix (P_{ij}) is the inverse of the intercorrelation matrix (r_{ij}). Subsidiary calculations lead to the regression coefficients and their standard errors. Also, if any of the coefficients, so calculated, are not significantly different from zero, the regression coefficients can be adjusted for the elimination of these variables. Two-tailed student's-t tests and 95 per cent confidence limits are used to determine the significance of these regression coefficients.

It was decided to compute a demand model for the years 1936 to 1952, and to use this model to predict the demand for transit for the years 1952 to 1957. A comparison of actual demand with predicted demand will permit the rating of these multiple regression models as prediction tools. Even though this rating is based only on one small sample, it will be a fair indication of their relative value.

After the development of Model I indicated that the form selected was sufficiently accurate, and that all of the regression coefficients were highly significant (5 per cent level), the next step was to attempt to deflate the importance of the time factors, by introducing other factors into the function explicitly. There are two very important requirements which the demand model must meet:

1. There must be sufficient variables in the function to permit it to adequately represent the "real world."
2. There must be sufficient variables in the function to permit management to evaluate the probable effects of changes in operating policy on demand for transit.

To fulfill these requirements, and at the same time, to test the power of the multiple regression method of model building, five new possible variables were used to develop Model II in addition to the four used in Model I. Each new variable was selected to represent some phase of the economic system. The nine variables, a rather large number for multiple regression methods, make it difficult to derive a sufficiently accurate function with as few observations as are available (12 observations for Model II). There are only three degrees of freedom (N-n) for the significance tests (the means are assumed to be zero). In spite of this difficulty, it was decided to force the method to do its utmost.

The first four variables used for Model II are, again, logarithms of per capita demand and adjusted average fare, time, and time squared. The fifth variable, logarithms of United States Index of Industrial Production (1935-1939 base), was used to represent the effects of changes in the national economic situation. The sixth variable, logarithms of registered passenger automobiles in Metropolitan Atlanta, was used to represent the effects of competition by other means of transportation. The seventh variable, logarithms of transit vehicle miles per capita served, was used to indicate the effects of the quantity of service offered by the company, and to place in the model another variable over which management has control (the other variable which management controls is the average fare). The eighth variable, logarithms of adjusted Effective Buying Income per Capita for Atlanta, was used to represent the effects of the size of individual incomes of prospective transit riders. The ninth variable, logarithms of adjusted Bank Clearings for Atlanta, was used to reflect the effects of the local economic conditions.[29]

These nine variables were normalized and the intercorrelation matrix was calculated. The inverse of this matrix was then determined, and an attempt was made to test its accuracy. All of this work was done using eight digit decimals on a standard desk calculator. Even though the

---

[29] See Appendix, pp. 72-76.

standard "check column" was used in the inversion procedure, the inverse was far from being accurate. Iterative methods are available for improving the accuracy of an inverse.[30] However, the size of this matrix, nine by nine, makes it impractical to proceed with hand computation. The round off errors accumulate to such a degree as to make it almost impossible to get sufficient accuracy. The variables for this model, and all subsequent models, were then used in conjunction with an I.B.M. 650 electronic computer to provide sufficient accuracy of calculation. A four part multiple regression subroutine which computes the intercorrelation matrix, its inverse, the regression coefficients, and their standard errors was used.[31] This subroutine was used to produce Model II in the following form:

$$X = b_o y^{b_1} e^{b_2 t + b_3 t^2} P^{b_4} A^{b_5} V^{b_6} R^{b_7} C^{b_8} \qquad (17)$$

where X is the transit rides demanded per capita served,

---

[30]Hotelling, Harold, "Some New Methods in Matrix Calculation," The Annals of Mathematical Statistics, Vol. XIV, 1943, pp. 14-16.

The Hotelling Iterative Method is as follows:

1. Let C_o be the trial inverse
2. Compute $C_{k+1} = C_k(2 - rC_k)$
3. Check on convergence: Is $rC_k = I$ to the desired number of places for each element of I?

where r is the intercorrelation matrix, and I is the unit matrix.

[31]International Business Machines' Multiple Regression Routine ST01. Phases I-IV, Rich Electronic Computer Center.

y is the adjusted average fare, t is time, t^2 is time squared, P is the U. S. Index of Industrial Production, A is the registered passenger automobiles for Metropolitan Atlanta, V is the transit vehicle miles per capita served, R is the adjusted effective buying income per capita, and C is the adjusted bank clearings for Atlanta.

After computation of all of the foregoing regression coefficients and their standard errors, student's-t tests were used to test the following null hypothesis:

$$H_o: \quad \beta_i = 0 \tag{18}$$

against the alternate hypothesis:

$$H_1: \quad \beta_i \neq 0 \tag{19}$$

The tests for each coefficient were calculated as follows:

$$t_i = \frac{b_i - \beta_i}{S_{b_i}} \quad \begin{cases} \text{N-n degrees of freedom} \\ i = 2, \ldots, 9 \end{cases} \tag{20}$$

where b_i is the calculated coefficient, β_i is its true value, S_{b_i} is the unbiased standard error of b_i, N is the number of observations, and n is the number of variables. The null hypothesis is to be rejected if t_i is greater than $+t_{N-n}$ and also if t_i is less than $-t_{N-n}$ with a probability of 0.05; otherwise the null hypothesis is to be accepted. A rejection is equivalent to a probability of 0.95 that the appropriate regression coefficient could not have occurred due to chance.

When the regression coefficient is not significantly different from zero, it should be eliminated, and a new model calculated. However, since the variables used for Model II did not reach significance at the five per cent level, that criterion was not strictly adhered to. Instead, for Model II the four least significant variables were eliminated without reference to any one particular significance level, and the four most significant were retained. The cause of this difficulty is that there are not sufficient observations, and there are too many variables to allow sufficient degrees of freedom. In this case N-n = 12 - 9 = 3 d.f. under which the null hypothesis is accepted for all eight independent variables at the five per cent level of significance in Model II (the means are assumed to be zero).

The coefficient of multiple correlation was calculated by using equation (21)

$$R = \sqrt{1 - \frac{1}{a_{11}}} \qquad (21)$$

where a_{11} is the first element in the first row of the inverse matrix. This coefficient (R) was then used in equation (22) to calculate the unbiased standard error of the regression estimate. (σ_1 is the standard deviation of the dependent variable).

$$S = \sigma_1 \sqrt{\frac{N}{N-n}} \sqrt{1 - R^2} \qquad (22)$$

or

$$S = \sigma_1 \frac{1}{\sqrt{a_{11}}} \sqrt{\frac{N}{N-n}}$$

To establish the validity of the coefficient of multiple correlation, which is a measure of "goodness of fit" of the surface, a student's-t test was applied as in equation (24).

$$t = \frac{R - R_o}{\sqrt{\frac{1 - R^2}{N - n}}} \quad \text{(N-n degrees of freedom)} \qquad (24)$$

where R is the calculated coefficient of multiple correlation, R_o is its true value, N is the number of annual observations, and n is the number of variables in the model (including the dependent variable). This test indicates whether the coefficient of multiple correlation calculated could have occurred due to chance at the five per cent level of confidence under the null hypothesis, $H_o : R_o = 0$, and the alternate hypothesis, $H_1 : R_o \neq 0$.

Model III was developed, with the aid of the computer, using data for the 1936-1952 period and the five variables; logarithms of transit demand per capita served, logarithms of adjusted average fare, time squared, logarithms of transit vehicle miles per capita served, and logarithms of adjusted effective buying income per capita. The other four variables (least significant according to the t-test) were discarded in the development of this model. Model III, again a multiplicative model, was derived in the following form:

$$X = b_o y^{b_1} e^{b_3 t^2} V^{b_6} R^{b_7} \tag{25}$$

where letter symbols identify the same variables as for Model II.[32] Once again this procedure included calculation of the regression coefficients and their standard errors, the coefficient of multiple correlation, related t-tests, and the unbiased standard error of the regression estimates.

With the development of Models II and III, the original objective was attained; but it was decided to also construct similar models (IV and V) for the entire 1936-1957 period (with the exception of the years 1942-1946). These models provide up-to-date decision tools for prospective use now and for a limited number of years in the future. Model IV was developed in the same manner as Model II, and Model V in the manner of Model III. The primary purpose of these models is not mere prediction of demand, but also the evaluation of changes in operating policy (e.g., price and volume of service) in terms of their probable effect on demand.

Coefficients of Elasticity of Demand

When the demand function is known, the elasticity of demand can be easily obtained. When the quantity demanded is a function of a single variable, e.g., price, the elastic-

---

[32]Supra, pp. 36-37.

ity of demand is defined as the ratio of the relative change in quantity demanded to the relative change in price, when the relative changes are infinitesimal.[33] In mathematical symbols,

$$\eta_{xy} = \frac{\frac{dx}{x}}{\frac{dy}{y}} = \frac{d\log x}{d\log y} = \frac{dx}{dy} \cdot \frac{y}{x} . \qquad (26)$$

If the quantity demanded is a function of more than one variable one must use the concept of partial elasticity of demand.[34] If the demand function is linear (e.g., $x = a + by + ct$) the partial elasticity is as in equation (27).

$$\eta_{xy.t} = \frac{\partial x}{\partial y} \cdot \frac{y}{x} \qquad (27)$$

The subscripts of η to the right of the dot are held constant, while those to the left of the dot are allowed to vary. When the coefficient of elasticity, η, is equal to unity in absolute value, the demand is neither elastic nor inelastic --the same amount will be spent regardless of price. When η is greater than unity, the demand is elastic and a small change in price causes a larger change in demand; when it is less than unity, the demand is inelastic,

[33]Schultz, op. cit., p. 190.

[34]Moore, H. L., Synthetic Economics (New York, 1929), p. 55.

and a small change in price causes a smaller change in demand.

One advantage of the multiplicative model, which has been selected as the functional form in this thesis, is the ease with which this form yields elasticities of demand. Another important advantage, is that the multiplicative (or logarithmic) form expresses the rates of shift of the demand curves in relative terms, which are independent of the units in which the commodity is measured.

Thus, if the demand function in question is Model II, equation (17), the partial elasticity is as in equation (28),

$$\eta_{xy.tt^2PAVRC} = \frac{\partial x}{\partial y} \cdot \frac{1}{x} = b_1 \tag{28}$$

where t, t^2, P, A, V, R, C are held constant, and X and y are allowed to vary. Also, the rate of shift is as follows:

$$\frac{1}{x} \cdot \frac{\partial x}{\partial t} = b_2 + 2(b_3)t. \tag{29}$$

The rate of shift of the demand curve is, therefore, a function of time, increasing or decreasing with time according to the sign of the second term of equation (29).

Elasticities of demand and rates of shift were calculated and analyzed for each model in an attempt to accurately describe the demand for transit in Atlanta.

Evaluation of Decision Models

<u>Prediction accuracy</u>.--To evaluate the probable effects of

changes in operating policy with the use of the demand models, it is necessary to translate management's decisions into adjusted average fare and/or transit vehicle miles per capita served. The predicted effect on demand can then be calculated by substituting these contemplated values along with the values for the other variables (secured by extrapolation) into the demand models.

To evaluate the prediction accuracy of multiple regression models in general, and to specifically evaluate Models II and III, two methods are possible. Both methods were used here to evaluate predicted transit demand for the years 1953-1957. The first method, which has as its aim the prediction of demand and not the evaluation of management decisions, involves the extrapolation of the regression estimates of demand for the years 1936-1952 into the 1953-1957 period. The second method involves the extrapolation of each variable into the 1953-1957 period by extending the smooth curve fitted to the 1936-1952 period. This method, which is similar to the procedure used in the evaluation of the probable effect of management's decisions on demand, should be a good test of the validity of the models for that purpose. Each method was evaluated by calculating the correlation between predicted and actual demand for the five year period, using both Model II and Model III. The average per cent error was also calculated to emphasize the relatively small magnitude of prediction error.

Goodness of fit.--Each model was designed to represent "the real world." The fit of the calculated demand surface to actual demand is an indication of this ability. Coefficients of multiple correlation, which apply to the period underlying the model, were used along with their significance tests to measure the goodness of fit. An analysis of the error size and distribution was also made to more fully evaluate the accuracy of the multiple regression models during the period for which they were constructed.

CHAPTER III

RESULTS

Presentation of Models

In the development of a tool to permit evaluation of management decisions at the company level, the Atlanta Transit System, Incorporated was chosen as the subject. This choice of subject was based on several factors of which convenience of communications, cooperation of management, interests of the writer, and prospective value to the subject were paramount. The results of this investigation follow.

The initial hand computation of Model I, which was an attempt to explore the importance of the time factors, resulted in the following dynamic demand model:

$$\log x = 3.2537 - 0.8440 \log y + 0.0242\,Mt - 0.0012\,Mt^2 \qquad (29)$$

$$\begin{array}{c} (0.1278) \quad (0.0001) \quad (0.0000003) \\ X = 1794.7721\, y^{-0.8840}\, e^{0.0242t - 0.0012t^2} \end{array} \qquad (30)$$

where X is the transit rides per capita, y is the adjusted average fare, t is time, and t^2 is time squared. The values in parentheses are the standard errors of the regression coefficients.

A summary of the descriptive statistics for Model I and its parameters are presented in Table 1.

Table 1. Summary of Descriptive Statistics for Model I (1936-1952)

Variable	Regression Coefficient	Standard Errors of Regression Coefficients	Student's "t" (8 d.f.)	Level of Significance (%)
2.Adj.Avg. Fare	-0.8840	0.1278	-6.92	5
3.Time	0.0242	0.0001	242.0	5
4.Time Squared	-0.0012	0.0000003	-4000.0	5

Coefficient of Multiple Correlation	= 0.987674
"Student's-t value" for Coefficient of Multiple Correlation (8 d.f., significant at 5% level)	= 19.9
Maximum value of Coefficient of Multiple Correlation which could have occurred due to chance (8 d.f. significant at 5% level)	= 0.811000
Unbiased Standard Error of the Regression Estimate	= 0.014202
Two Standard Errors	= 0.028404
Three Standard Errors	= 0.042606

It is noted that the standard error of the regression estimates in Table 1 apply only to the logarithmic form of the model, equation (29). If it is assumed that the distribution of estimates is normal, it can be said that 99.73 per cent of the actual values will have values within (plus or minus three standard errors) 0.042606 of the estimated values of demand. Even though the assumption of normality might be void, by Tchebysheff's inequality an area under any distribution of plus or minus three standard errors (when σ is known) includes 90 per cent of the values. As a result, three standard errors provide confidence limits for the range of estimates.

The coefficient of multiple correlation of 0.987 (being close to one) is indicative of the "goodness of fit" between theoretical and predicted values. Interpretation of this correlation should consider that there are only twelve observations available in four dimensions (four variables) with which one would expect the coefficient to be high. Since all the regression coefficients are highly significant at the five per cent level, the form of this function is judged to be accurate.

As was stated in the procedures, five new variables were introduced to deflate the effect of the time factors.[35] The computation of Model II was performed on an IBM 650 electronic computer. The model is as follows:

$$\begin{aligned} \log x = {} & 7.2214 - 0.5201 \log y + 0.0359\, Mt \\ & - 0.0012\, Mt^2 + 0.5459 \log P \\ & - 0.0283 \log A + 0.8552 \log V \\ & + 0.4510 \log R - 0.8528 \log C \end{aligned} \qquad (31)$$

or

$$X = 16651150\, y^{\overset{(0.5759)}{-0.5201}}\, e^{\overset{(0.0645)}{0.0359t} - \overset{(0.0009)}{0.0012t^2}}\, P^{\overset{(1.1834)}{0.5459}}\, A^{\overset{(0.7472)}{-0.0283}}\, V^{\overset{(0.6682)}{0.8552}}\, R^{\overset{(0.4928)}{0.4510}}\, C^{\overset{(2.2191)}{-0.8528}} \qquad (32)$$

[35]Supra, p. 35.

where x is the per capita demand, y is the adjusted average fare, t is time, t^2 is time squared, P is the U. S. Index of Industrial Production, A is registered passenger automobiles, V is transit vehicle miles per capita served, R is adjusted effective buying income per capita, and C is adjusted bank clearings. The figures in parentheses are, again, the standard errors of the regression coefficients. A summary of descriptive statistics for Model II are presented in Table 2.

Table 2. Summary of Descriptive Statistics for Model II (1936-1952)

Variable	Regression Coefficient	Standard Errors of Regression Coefficients	Student's "t" (3 d.f.)	Level of Significance (%)
2 Adj.Avg.Fare	-0.5201	0.5759	-0.9031	50
3 Time	0.0359	0.0645	0.5561	70
4 Time Squared	-0.0012	0.0009	-1.3166	30
5 US Ind.of I.P.	0.5459	1.1834	0.4613	70
6 Reg. Autos	-0.0283	0.7472	-0.0379	--
7 Vehicle Miles	0.8552	0.6680	1.2803	30
8 Effect. Inc.	0.4510	0.4928	0.9151	50
9 Bank Clear.	-0.8528	2.2191	-0.3843	80

Coefficient of Multiple Correlation = 0.994935
"Student's-t value" for Coefficient of Multiple Correlation (3 d.f., significant at 5% level) = 12.29

Unbiased Standard Error of the Regression Estimate = 0.018238
Two Standard Errors = 0.036476
Three Standard Errors = 0.054714

Since this model has nine variables and only twelve observations, there are only three degrees of freedom available. With this restriction it is difficult to secure significant t-values for the regression coefficients. Each significance level reported in Table 2 is that at which the coefficient is barely significant. The coefficient of multiple correlation, 0.994935, was highly significant at the five per cent level, indicating a very good fit for the demand surface.

The standard error of the estimates, 0.018238, is only slightly higher than that for Model I (a difference of 0.004036). It can be said, then, that at least 90 per cent of these estimates will have values within 0.054714 of the actual demand (plus or minus three standard errors).

The five per cent level of significance, selected in this study as appropriate for rejection of the null hypothesis that the regression coefficients are equal to zero, cannot be used for this model.[36] The four variables of highest significance, adjusted average fare, time squared, transit vehicle miles per capita served, and adjusted effective buying income per capita, were retained for the construction of Model III. The other four variables were discarded. The necessity for relaxation of the acceptable significance level has been encountered previously in the Econometrics field. As Professor Schultz states:[37]

---

[36]Supra, pp. 37-38.

[37]Schultz, op. cit., p. 214.

> . . . We must supplement our purely statistical tests of significance with all the theoretical and factual knowledge at our disposal. . . .
>
> There is another and even more compelling reason why our standard errors give us little aid in drawing probable inferences: They are all derived from time series. Now time series, especially those relating to social and economic phenomena, are likely to violate in a marked degree the fundamental assumption which underlies the use of the methods sketched above (significance tests), namely, that not only the successive items in the series but also the successive parts into which the series may be divided must be random selections from the <u>same</u> universe.

Model III was then constructed using only five variables. The electronic computation techniques lead to the following demand model:

$$\log x = -1.2033 - 0.2410 \log y - 0.0001\ Mt^2 + 1.7251 \log V + 0.3039 \log R \qquad (33)$$

or

$$X = 0.1597\ y^{\overset{(0.1493)}{-0.2410}}\ e^{\overset{(0.0001)}{-0.0001t^2}}\ V^{\overset{(0.3227)}{1.7251}}\ R^{\overset{(0.1453)}{0.3039}} \qquad (34)$$

where X is transit rides per capita served, t^2 is time squared, V is transit vehicle miles per capita served, R is adjusted effective buying income per capita served. A summary of descriptive statistics for Model III are reported in Table 3.

The standard error of Model III, 0.017527, is 0.000711 less than that for Model II. The coefficient of multiple correlation, 0.989066, is not as high as that for Model II, however, it is more reliable since seven degrees of freedom are now available. Each of the four variables, except time squared, has better significance than they had in the previous model; but the variable transit vehicle miles per capita served is the only one to reach significance at the five per cent level. Since the inclusion of additional variables was intended to deflate the effect of the time factors, it is encouraging to note that time squared is quite insignificant.

Table 3. Summary of Descriptive Statistics for Model III (1936-1952)

Variable	Regression Coefficient	Standard Errors of Regression Coefficients	Student's "t" (7 d.f.)	Level of Significance(%)
2 Adj.Avg.Fare	-0.2410	0.1493	-1.6139	20
4 Time Squared	-0.0001	0.0001	-0.8361	50
7 Vehicle Miles	1.7251	0.3227	5.3458	5
8 Effect. Inc.	0.3039	0.1453	2.09126	10

Coefficient of Multiple Correlation	= 0.989066
"Student's-t value" for Coefficient of Multiple Correlation (7 d.f., significant at 5% level)	= 18.6
Maximum value of Coefficient of Multiple Correlation which could have occurred due to chance (7 d.f., significant at 5% level)	= 0.838000
Unbiased Standard Error of the Regression Estimate	= 0.017527
Two Standard Errors	= 0.035054
Three Standard Errors	= 0.052581

In a similar manner Models IV and V were developed for the period 1936-1957.[38] Model IV, with nine variables is as follows:

[38] Supra, pp. 34-40.

$$\log x = -3.1140 - 0.4239 \log y + 0.0053\ Mt - 0.0009\ Mt^2 - 0.0285 \log P + 0.1303 \log A + 0.6796 \log V + 0.4165 \log R + 0.2909 \log C \qquad (35)$$

or

$$X = 0.0013\ y^{\overset{(0.1739)}{-0.4239}}\ e^{\overset{(0.0160)}{0.0053t}\ -\ \overset{(0.0004)}{0.0009t^2}}\ P^{\overset{(0.1505)}{-0.0285}}\ A^{\overset{(0.2953)}{0.1303}}\ V^{\overset{(0.2876)}{0.6796}}\ R^{\overset{(0.1770)}{0.4165}}\ C^{\overset{(0.2221)}{0.2909}} \qquad (36)$$

where letter symbols represent the same variables as for Model II.[39] A summary of descriptive statistics for Model IV is reported in Table 4.

Comparison of Model IV with Model II showed that although the standard error of Model IV is larger (a difference of 0.025279) than that for Model II, the coefficient of correlation for Model IV is slightly higher. This high order of correlation and its associated eight degrees of freedom led the writer to rate Model IV as being more statistically reliable than Models I, II and III.

The results of the significance tests for the coefficients in Model IV emphasize a point made earlier. The five additional observations, and five extra degrees of freedom, do much to aid in the development of a demand model

---

[39] *Supra*, p. 48.

whose coefficients are significantly different from zero. The four variables, adjusted average fare, time squared, vehicle miles per capita served, and adjusted effective buying income per capita, succeeded in reaching the five per cent level of significance. These are the same four variables which were retained in the development of Model III.

Table 4. Summary of Descriptive Statistics for Model IV (1936-1957)

Variable	Regression Coefficient	Standard Error of Regression Coefficient	Student's "T" (8 d.f.)	Level of Significance(%)
2 Adj.Avg.Fare	-0.4239	0.1739	-2.4379	5
3 Time	0.0053	0.0160	0.3299	80
4 Time Squared	-0.0009	0.0004	-2.5584	5
5 US Ind. I.P.	-0.0285	0.1505	-0.1892	90
6 Reg.Pass.Ats.	0.1303	0.2953	0.4410	70
7 Tran.Veh.Mi.	0.6796	0.2876	2.3628	5
8 Adj.Eff.Inc.	0.4165	0.1770	2.3533	5
9 Bank Clear.	0.2909	0.2221	1.3099	30

Coefficient of Multiple Correlation = 0.995790
"Student's-t value" for Coefficient of Multiple Correlation (8 d.f., significant at 5% level) = 20.05

Unbiased Standard Error of the Regression Estimate = 0.043517
Two Standard Errors = 0.087034
Three Standard Errors = 0.130551

After elimination of the less significant factors, Model V was developed in the following form:

$$\log x = -0.5979 - 0.3324 \log y - 0.00003\, Mt^2 + 1.2914 \log V + 0.3624 \log R \qquad (37)$$

or

$$X = 0.2524\, y^{\overset{(0.1365)}{-0.3324}}\; e^{\overset{(0.00007)}{-0.00003t^2}}\; V^{\overset{(0.2122)}{1.2914}}\; R^{\overset{(0.1323)}{0.3624}} \tag{38}$$

where the letter symbols refer to the same variables as for Model III.[40] A summary of descriptive statistics for Model V are reported in Table 5.

Table 5. Summary of Descriptive Statistics for Model V (1936-1957)

Variable	Regression Coefficient	Standard Error of Regression Coefficient	Student "t" (12 d.f.)	Level of Significance(%)
2 Adj.Avg.Fare	-0.3324	0.1365	-2.4359	5
4 Time Squared	-0.00003	0.00007	-0.4056	70
7 Trans.Veh.Mi.	1.2914	0.2122	6.0858	5
8 Adj.Eff.Inc.	0.3624	0.1323	2.7396	5

Coefficient of Multiple Correlation	= 0.990111
"Student's-t value" for Coefficient of Multiple Correlation (12d.f., significant at 5% level)	= 24.4
Maximum value of Coefficient of Multiple Correlation which could have occurred due to chance (12 d.f., significant at 5% level)	= 0.722000
Unbiased Std. Error of Regression Estimate	= 0.027558
Two Standard Errors	= 0.055116
Three Standard Errors	= 0.082674

The standard error of Model V is not as large as that for Model IV, but is larger than that for Model III (a difference of 0.010031). The coefficient of multiple correlation,

[40] Supra, p.50.

0.990111, is slightly lower than those for the other models. This is to be expected, since there are twelve degrees of freedom available for Model V. It can be seen in Table 5, that all variables, except time squared, are significant at the five per cent level. Thus, the effect of the time variable has been successfully reduced by the inclusion of additional variables which affect the demand for transit.

Coefficients of Elasticity of Demand[41]

It is possible to speak of an elasticity of demand for each variable in the demand function, i.e., price elasticity of consumption (demand), service (vehicle miles per capita) elasticity of consumption, etc. Each elasticity is taken directly from the exponents of the variables, when the models are in multiplicative form. Also, the standard error of the coefficient of elasticity of demand is the same as the standard error of the regression coefficient when the model is in this form. Since this research is primarily concerned with the evaluation of management decisions in terms of their effects on demand, only those elasticities pertinent to the variables under management's control will be analyzed. Table 6 is a summary of the elasticities of demand yielded by each model for the variables, adjusted average fare (real price), and transit vehicle miles per capita served (service).

It can be seen that these elasticities as well as their standard errors vary with different models. Models III

---

[41] Supra, p. 40.

and V, the five variable models, are considered by the writer to be the most desirable of the five models, in the light of their descriptive statistics and his overall knowledge of the dynamics of transit demand. Therefore, the following statements apply to these two models only.

Table 6. Elasticities of Demand

Model	Price Elasticity	Standard Error of Price Elasticity	Service Elasticity	Standard Error of Service Elasticity
I (1936-1952)	-0.8840	0.1278	-	-
II (1936-1952)	-0.5201	0.5759	0.8552	0.6682
III (1936-1952)	-0.2410	0.1493	1.7251	0.3227
IV (1936-1957)	-0.4239	0.1739	0.6796	0.2876
V (1936-1957)	-0.3324	0.1365	1.2914	0.2122

The price elasticity of demand for Model III is interpreted to mean that, if the adjusted average fare which prevailed in any year between 1936 and 1952 had been decreased (or increased) by one per cent and <u>if the demand surface had remained fixed for one year</u>, there would have been an increase (or decrease) of approximately 0.24 of one per cent in the annual per capita demand for transit. Similarly, the price elasticity of demand for Model V means that, if the adjusted average fare which prevailed in any year between 1936 and 1957 had been decreased (or increased) by one per cent and if the demand surface had remained fixed for one year, there would have been an increase (or decrease) of approximately 0.33 of one per cent in the annual per capita demand for transit.

The service elasticity of demand for Model III is interpreted to mean that, if the amount of service offered by the transit company (transit vehicle miles per capita served) during any year between 1936 and 1952 had been increased (or decreased) by one per cent and if the demand surface had remained fixed for one year, there would have been an increase (or decrease) of approximately 1.73 per cent in annual per capita demand for transit. For Model V the service elasticity means that, if the amount of service offered by the transit company during any year between 1936 and 1957 had been increased (or decreased) by one per cent and if the demand surface had remained fixed for one year, there would have been an increase (or decrease) of approximately 1.29 per cent in annual per capita demand for transit.

These elasticities of demand are independent of the rate of shift of the demand surface.[42] Nevertheless, the underlined conditions in the foregoing statements are necessary in order to emphasize that it takes time for demand to adjust to changes in these control variables. When the

---

[42] Schultz, *op. cit.*, p. 198.

demand surface is derived from annual data, the necessary time interval could very well be as great as one year.

The rates of shift of the demand surfaces are derived, similarly to elasticities, directly from the demand function.[43] The rates of shift for the five models are reported in Table 7.

Table 7. Rates of Shift of the Demand Surfaces

Model	Rate of Shift of Demand Curve	Rate of Shift in 1957
I (1936-1952)	0.0242 - 0.0024 t (0.0001) (0.0000003)	-0.0770
II (1936-1952)	0.0359 - 0.0024 t (0.0645) (0.0009)	-0.0887
III (1936-1952)	-0.0002 t (0.0001)	-0.0044
IV (1936-1957)	0.0053 - 0.0018 t (0.0160) (0.0004)	-0.0449
V (1936-1957)	-0.00006 t (0.00007)	-0.0020

The rate of shift of the demand surface for Model III means that, even if all the independent variables in the function other than time remained fixed, the per capita demand for transit during the period 1936-1952 would have decreased approximately at the rate of 0.0002(t) of one per cent per annum. The rate of shift is seen to be a function

[43] *Supra*, p. 42.

of time, the rate growing as time passes. The current rate of shift is seen to be downward at approximately 0.0044 of one per cent per annum, even if the other factors remain constant.

The rate of shift for Model V, similarly, means that even if all factors other than time would have remained fixed, the per capita demand would have decreased by approximately 0.0020 of one per cent per annum.

Evaluation of Decision Models

Prediction accuracy.--The prediction accuracy of Models II and III (1936-1952), when predicting demand for the year 1953 through 1957, was evaluated by two methods. The first was by extrapolation of the estimates of demand for each model. The second was by extrapolation of the individual variables with calculation of the regression estimates of demand using each model. The correlation between estimated and observed demand was used to rate the accuracy of prediction. Table 8 reports these results.

All of the correlations are high, and for all practical purposes are equal. However, Model II has consistently higher correlation, and the method which involves extrapolation of each variable is also consistently better than the method of extrapolating the regression estimates. The figures in parentheses, the average per cent error in estimation, indicated, however, that the method of extrapolating estimated demand had a much smaller average per cent error

(0.25%) than the other method (4.1%). It can also be seen that Model III was more accurate than Model II. Model III had an average per cent error for the two methods of 0.83 per cent, while Model II had an average per cent error of 4.03 per cent.

Table 8: Correlation Between Estimated and Actual Demand

Model	Extrapolation Regression Estimates of Demand	Extrapolation of variables and calculation of Demand
II	0.995726	0.997123
Avg. per cent error	(0.25%)	(7.8%)
III	0.990717	0.996876
Avg. per cent error	(0.25%)	(1.4%)

The foregoing statistics indicate that Model III, the five variable model, is the more accurate decision tool for the period 1953 through 1957. Since Model V is analogous to Model III, but applicable to the period 1936-1957, the writer feels that the former model would be best for use at present and for a limited number of years in the future. Since the evaluation of management decisions is analogous to the method involving extrapolation of individual variables, the only available estimate of the probable average per cent error for evaluation purposes is 4.1 per cent.

Goodness of fit.--Each model was designed to represent "reality" by the inclusion of significant variables. The coefficients of multiple correlation and the average per cent error in the regression estimates are indicators of this ability.

Table 9. Goodness of Fit of Demand Surface

Model		Coefficient of Multiple Correlation	Average Per cent error in Regression Estimates (%)	Average Magnitude of Error (Passengers)
II	(1936-1952)	0.994935	1.91	1,515,100
III	(1936-1952)	0.989066	2.46	2,114,226
IV	(1936-1957)	0.995790	1.93	1,409,633
V	(1936-1957)	0.990111	2.73	2,083,765

From Table 9 it is evident that all the coefficients of multiple correlation show good fit of the demand surface, or good representation of reality for all the models. However, the average per cent error permits a more discriminating evaluation of the models. Even though the elimination of insignificant variables from Models II and IV was motivated by statistical considerations, it did in both cases slightly reduce the "goodness of fit" of the surface.

The average magnitude of error for each model is also reported in Table 9. Errors of nagnitude one to two million passengers are considered quite good when estimating an annual demand of the order of 110 million passengers.

Significant Factors

The foregoing results indicate that the four variables which most significantly affect the demand for transit are adjusted average fare, time squared, transit vehicle miles per capita served, and adjusted effective buying income per capita. Another indicator of their relative importance as affectors is their partial correlation coefficients. These are shown in Table 10.

Table 10. Partial Correlation Coefficients

Variable	Model	Partial Correlation Coefficient
Adjusted Avg. Fare	III	0.520748
	V	0.575212
Time Squared	III	0.310026
	V	0.116464
Transit Veh.Mi.per Cap.	III	0.896257
	V	0.869043
Adj. Effec. Buying Inc.	III	0.620101
	V	0.620314

These partial correlation coefficients are a measure of the linear relationship between each variable and the per capita demand which remains after any dependence on the remaining variables has been removed.[44] Thus, they are also a measure of the fraction of the total regression due to each variable. It is observed that the amount of service offered by the transit company plays the major role.

[44]Bennett and Franklin, Statistical Analysis in Chemistry and the Chemical Industry, New York, 1954, p. 287.

CHAPTER IV

CONCLUSIONS AND RECOMMENDATIONS

Conclusions.--The purposes of this study are several. The illustration of the econometrician's method of approach and its value to the industrial engineer is one of the most important purposes. Other purposes are: to show the applicability of demand theory at the company level, to evaluate the use of multiple regression models as decision tools, to investigate the dynamics of demand in a service industry, and to locate specific factors which affect transit service demand and formulate them in a dynamic demand model. The results obtained in this research permit the following conclusions:

1. The methods of the econometrician are of value to the industrial engineer, since they permit quantitative analysis of economic and industrial system parameters. The methods make use of historical data, which might not otherwise be used, to develop accurate mathematical functions.
2. The modification of the requirements underlying the assumptions regarding mass market phenomena, hypothesized to the enterprise level, does not seriously affect or curtail the accuracy of the method of approach or the demand functions so calculated.
3. The multiple regression models developed are accurate

representations of "reality." They are difficult to construct when the number of variables exceeds four. Machine computation, however, makes possible the development of functions which, for all practical purposes, are unlimited, except by the size of the electronic computer used.

4. The demand models developed are accurate decision tools for evaluation of the effects of changes in operating policy on the demand for transit service, when evaluated by the criterion of prediction for a sample of five years (1953-1957). No conclusion can be made concerning the accuracy of Models IV and V since time must evolve before their evaluation is possible by comparison methods.
5. The price elasticity of demand and the service elasticity of demand indicate that an increase in the amount of service offered will bring about a greater increase in demand than an equivalent decrease in price.
6. The rates of shift of the demand surfaces indicate a general down-trend in transit demand. As indicated by Model V this is due to increasing cost of transit to customers, decreasing service offered, and other underlying time-trends.
7. The four factors which most significantly affect the demand for transit service in order of importance are:

transit vehicle miles per capita served, adjusted effective buying income per capita, adjusted average fare, and time squared.

Recommendations.--During the course of this research difficulties arose because of the small number of observations available. It is recommended that a study using semi-annual data, which would double the observations, be made to alleviate the difficulties of small degrees of freedom. This type of analysis would yield a short time demand function which would reduce the waiting time between changes and their affects on demand. It would probably lead to the investigation of seasonal factors as well as weather effects; but if data is published they should not pose any insurmountable difficulties.

It is recommended that the demand functions developed be adapted to analog computer techniques to determine optimum price policy and/or optimum service policy at which it is most advantageous for the transit company to operate.

Other questions which were not experimentally answered are: How is the demand function affected by abnormal periods (e.g., the war years)? Would better results be obtained if the period was separated into two parts --one before and one after the war? What is an appropriate significance level for t-tests in demand theory?

A mathematical model of the type developed is only accurate for a limited number of years. An extension of this research could include the development of control chart techniques to indicate when model revision is necessary.

APPENDIX

Table 11. Data for Adjusting Variables for Population Growth and Inflation Trends

Year	Metropolitan Area Population[1]	Atlanta Population Served by Transit (000)[2]	Consumer Price Index[3]
1936	402,450(As of 1930)	329	58.8
37	408,450 "	333	60.7
38	413,000 "	337	58.9
39	420,600 "	341	58.3
40	442,294(As of 1940)	346	58.4
41	457,000 "	356	61.7
42	466,000 "	362	68.3
43	473,000 "	369	73.1
44	488,000 "	376	74.3
45	497,500 "	384	76.9
46	525,000 "	404	82.7
47	550,000 "	422	95.8
48	565,000 "	431	102.2
49	590,000 "	433	102.0
50	668,022(As of 1950)	436	103.5
51	678,000 "	440	113.3
52	695,000 "	445	116.0
53	725,000 "	450	117.1
54	744,000 "	457	116.8
55	808,853 "	479	116.4
56	862,000(As of 1956)	483	118.3
57	885,000 "	497	121.3

[1]Atlanta Chamber of Commerce, Industrial Bureau.

[2]Estimated by Atlanta Transit System Incorporated.

[3]U. S. Department of Labor, Bureau of Labor Statistics, Atlanta, Georgia Year Average, Series A-2 (1947-1949 = 100 base).

Table 12. Transit Demand Series --Independent Variable (X)

Year	Revenue Passengers[1]	Transit Rides Per Capita Served	Logarithms of Transit Rides per Capita Served
1936	48,371,691	147.0	2.16732
37	48,989,641	147.1	2.16761
38	49,850,682	147.9	2.16997
39	54,753,658	160.6	2.20575
40	59,270,002	171.3	2.23376
41	64,875,325	182.2	2.26055
42	85,956,044	237.4	2.37548
43	107,397,845	291.0	2.46389
44	111,949,416	297.7	2.47378
45	118,022,834	307.4	2.48770
46	122,901,727	304.2	2.48316
47	115,593,639	273.9	2.43664
48	110,482,505	256.3	2.40875
49	102,386,634	236.5	2.37383
50	94,183,391	216.0	2.33445
51	87,332,742	198.6	2.29798
52	81,816,351	183.9	2.26458
53	75,154,579	167.0	2.22272
54	69,560,930	152.2	2.18241
55	66,454,038	138.7	2.14208
56	62,263,097	128.9	2.11025
57	56,142,785	113.0	2.05308

[1]1942 through 1946 not included in analysis. Years 1946, 1949, 1950 corrected for strikes by Atlanta Transit System, Incorporated.

Period 1936-1952

Mean = 2.276783

Standard Deviation = 0.090724

Period 1936-1957

Mean = 2.237176

Standard Deviation = 0.102849

Table 13. Transit Price Series --(Y)

Year	Passenger Revenue($)[1]	Average Fare(¢)	Adjusted Average Fare (¢)	Logarithms of Adjusted Average Fare
1936	3,654,826	7.55	12.84	1.10857
37	3,693,694	7.53	12.41	1.09377
38	3,743,931	7.51	12.75	1.10551
39	4,043,194	7.38	12.66	1.10243
40	4,256,359	7.18	12.29	1.08955
41	4,652,878	7.17	11.62	1.06521
42	6,190,891	7.20	10.54	1.02284
43	7,816,302	7.27	9.95	0.99782
44	8,188,973	7.31	9.84	0.99300
45	8,574,067	7.26	9.44	0.97497
46	8,784,946	7.14	8.63	0.93601
47	8,419,605	7.28	7.60	0.88081
48	8,689,847	7.86	7.69	0.88593
49	8,710,347	8.50	8.33	0.92065
50	7,931,974	8.42	8.14	0.91062
51	9,305,055	10.65	9.40	0.97313
52	9,311,371	11.38	9.81	0.99167
53	9,243,780	12.29	10.50	1.02119
54	8,837,343	12.30	10.53	1.02243
55	8,852,099	13.32	11.44	1.05843
56	8,759,508	14.06	11.89	1.07518
57	8,860,125	15.78	13.01	1.11428

[1]1942 through 1946 not included in analysis.

<u>Period 1936-1952</u>

Mean = 1.003158

Standard Deviation = 0.084547

<u>Period 1936-1957</u>

Mean = 1.019376

Standard Deviation = 0.077686

Table 14. Time Series (t) and Time Squared Series (t^2)

Year	Time[1]	Time Squared[1]
1936	1	1
37	2	4
38	3	9
39	4	16
40	5	25
41	6	36
42	7	49
43	8	64
44	9	81
45	10	100
46	11	121
47	12	144
48	13	169
49	14	196
50	15	225
51	16	256
52	17	289
53	18	324
54	19	361
55	20	400
56	21	441
57	22	484

[1]1942 through 1946 not included in analysis.

Time Series: 1936-1952

Mean = 9.0

Standard Deviation = 5.759051

Time Squared Series: 36-52

Mean = 114.166666

Standard Deviation = 105.3809

Time Series: 1936-1957

Mean = 12.235294

Standard Deviation = 7.008645

Time Squared Series 36-57

Mean = 198.823529

Standard Deviation = 161.1869

Table 15. United States Index of Industrial Production (P)
The Effect of National Economic Situation

Year	Index[1]	Logarithms of Index
1936	104	2.01703
37	113	2.05308
38	87	1.93952
39	109	2.03743
40	126	2.10037
41	168	2.22531
42	212	-
43	258	-
44	252	-
45	214	-
46	177	-
47	194	2.28780
48	198	2.29667
49	183	2.26245
50	209	2.32015
51	229	2.35984
52	230	2.36173
53	252	2.40140
54	235	2.37107
55	261	2.41664
56	268	2.42813
57	273	2.43616

[1]United States Department of Commerce, Historical Statistics of the United States --1789-1945, 1946-1952. 1952-1957, (1935-1939 = 100 base). Years 1953-1957, reported in another base, were transformed into the 1935-1939 base by estimation. Years 1942 through 1946 not included in analysis.

Period 1936-1952

Mean = 2.188450

Standard Deviation = 0.143010

Period 1936-1957

Mean = 2.253812

Standard Deviation = 0.157624

Table 16. Registered Passenger Automobiles in Metropolitan Atlanta (A) --The Effect of Competition

Year	Registered Passenger Automobiles[1]	Logarithms of Registered Passenger Automobiles
1936	69,514	4.84207
37	89,361	4.95115
38	87,080	4.93992
39	95,756	4.98107
40	103,957	5.01685
41	98,704	4.99434
42	99,650	-
43	99,547	-
44	-	-
45	-	-
46	91,854	-
47	115,791	5.06100
48	129,694	5.11290
49	151,282	5.17978
50	161,342	5.20774
51	179,044	5.25295
52	193,332	5.28630
53	228,204	5.35832
54	242,869	5.38536
55	281,980	5.45022
56	282,668	5.45128
57	296,627	5.47222

[1]Atlanta Chamber of Commerce, Industrial Bureau. Years 1942 through 1946 not included in analysis.

<u>Period 1936-1952</u>

Mean = 5.068850

Standard Deviation = 0.133113

<u>Period 1936-1957</u>

Mean = 5.173153

Standard Deviation = 0.197939

Table 17. Transit Vehicle Miles Per Capita Served (V) -- The Effect of the Quantity of Service Offered

Year	Transit Vehicle Miles[1]	Transit Vehicle Miles Per Capita Served	Logarithms of Vehicle Miles Per Capita Served
1936	12,065,527	36.7	1.56467
37	12,584,442	37.8	1.57749
38	12,562,071	37.3	1.57171
39	12,979,372	38.1	1.58092
40	13,751,683	39.7	1.59879
41	14,572,021	40.9	1.61172
42	16,478,896	45.5	-
43	18,224,153	49.4	-
44	18,557,516	49.4	-
45	18,243,391	47.5	-
46	19,617,873	48.5	-
47	19,895,701	47.1	1.67302
48	19,529,288	45.3	1.65610
49	20,314,828	46.9	1.67117
50	19,008,831	43.6	1.63949
51	18,576,982	42.2	1.62531
52	17,703,388	39.8	1.59988
53	17,016,422	37.8	1.57749
54	15,621,669	34.2	1.53403
55	15,258,829	31.9	1.50379
56	14,865,844	30.8	1.48855
57	14,183,063	28.5	1.45484

[1]Years 1936-1938 include Atlanta Coach Company. Years 1942 through 1946 not included in analysis. Years 1946, 1949, 1950 were adjusted for strikes by Atlanta Transit System, Incorporated.

<u>Period 1936-1952</u>

Mean = 1.614192

Standard Deviation = 0.036952

<u>Period 1936-1957</u>

Mean = 1.584059

Standard Deviation = 0.060436

Table 18. Effective Buying Income per Capita of Atlanta (R) --The Effect of the Level of Individual Incomes

Year	Effective Buying Income($)[1] (000)	Adjusted Effective Buying Income per Capita	Logarithms of Effective Buying Income Per Capita
1936	324,695	1372.10	3.13738
37	260,094	1049.07	3.02080
38	251,655	1034.52	3.01471
39	288,548	1176.74	3.07067
40	296,612	1148.33	3.06007
41	353,989	1273.15	3.10490
42	464,095	-	-
43	673,809	-	-
44	654,493	-	-
45	611,208	-	-
46	690,470	-	-
47	796,714	1512.08	3.17958
48	844,589	1462.67	3.16515
49	867,553	1441.60	3.15885
50	912,787	1320.20	3.12065
51	1,004,839	1308.09	3.11664
52	1,177,211	1460.20	3.16441
53	1,237,262	1457.36	3.16950
54	1,321,481	1520.71	3.18204
55	1,363,091	1447.78	3.16071
56	1,574,793	1544.30	3.18873
57	1,587,044	1478.38	3.16979

[1]Atlanta Chamber of Commerce, Industrial Bureau. Years 1942 through 1946 not included in analysis.

<u>Period 1936-1952</u>

Mean = 3.109483

Standard Deviation= 0.054187

<u>Period 1936-1957</u>

Mean = 3.128500

Standard Deviation = 0.054495

Table 19. Bank Clearings in Atlanta (C) --The Effect of the Local Economic Situation

Year	Bank Clearings[1] (000)($)	Adjusted Bank Clearings(000)	Logarithms of Adjusted Bank Clearings
1936	2,601,000	4,423,469	9.64582
37	2,879,900	4,744,481	9.67624
38	2,670,765	4,534,405	9.65648
39	3,009,375	5,161,878	9.71282
40	3,430,900	5,874,829	9.76901
41	4,550,900	7,375,851	9.86782
42	5,467,454	-	-
43	6,560,573	-	-
44	7,629,200	-	-
45	8,263,900	-	-
46	9,886,579	-	-
47	10,803,900	11,277,557	10.05123
48	11,886,000	11,630,137	10.06558
49	11,573,153	11,346,228	10.05484
50	12,910,100	12,473,527	10.09601
51	15,178,383	13,396,631	10.12700
52	16,234,600	13,995,345	10.14598
53	16,433,100	14,033,390	10.14715
54	16,597,500	14,210,188	10.15259
55	18,597,100	15,976,890	10.20350
56	19,622,362	16,586,950	10.29274
57	20,556,000	16,946,413	10.31294

[1]Atlanta Chamber of Commerce, Industrial Bureau. Years 1942 through 1946 not included in analysis.

<u>Period 1936-1952</u>

Mean = 9.905833

Standard Deviation = 0.193936

<u>Period 1936-1957</u>

Mean = 9.998824

Standard Deviation = 0.220715

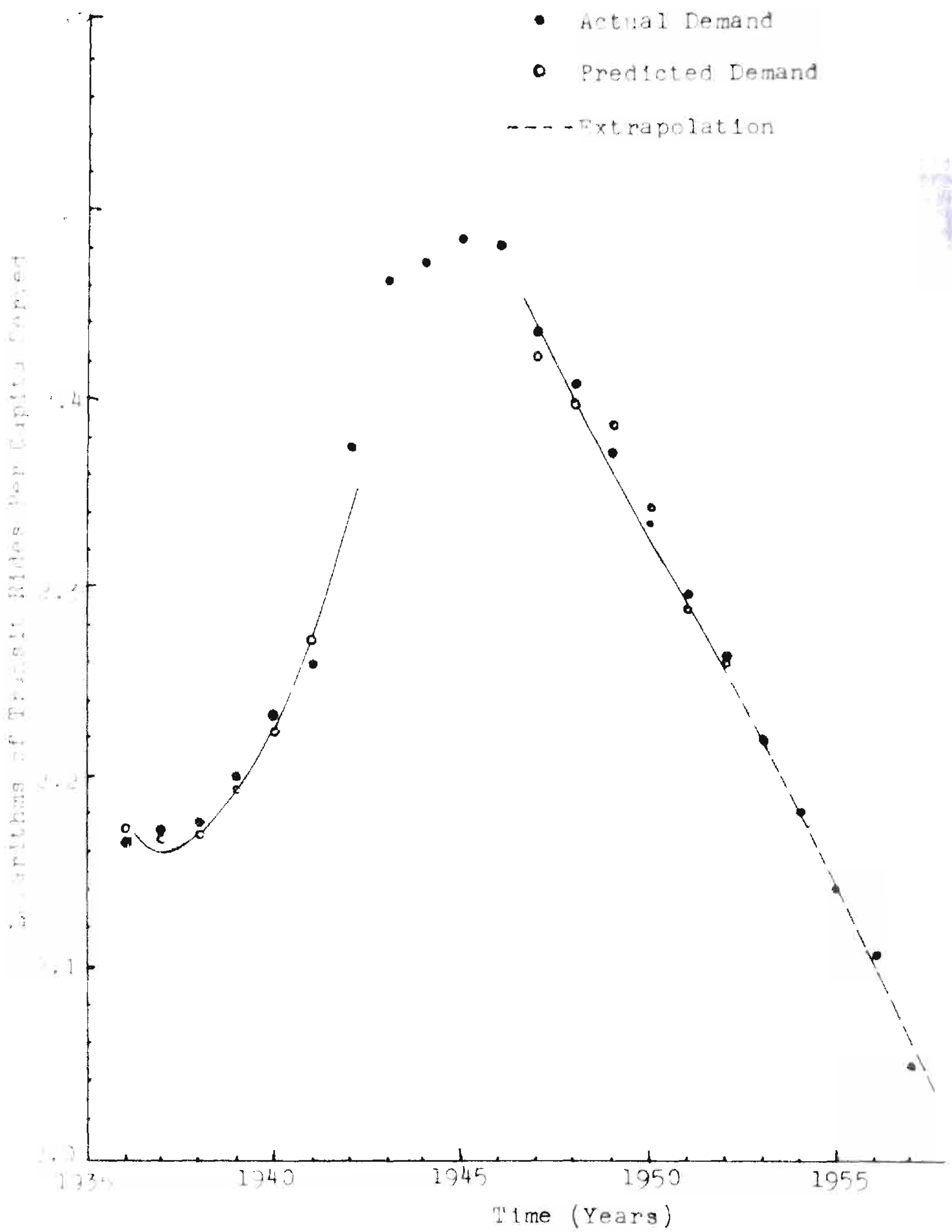


Fig. 6 : Extrapolation of Predicted Demand - Model II

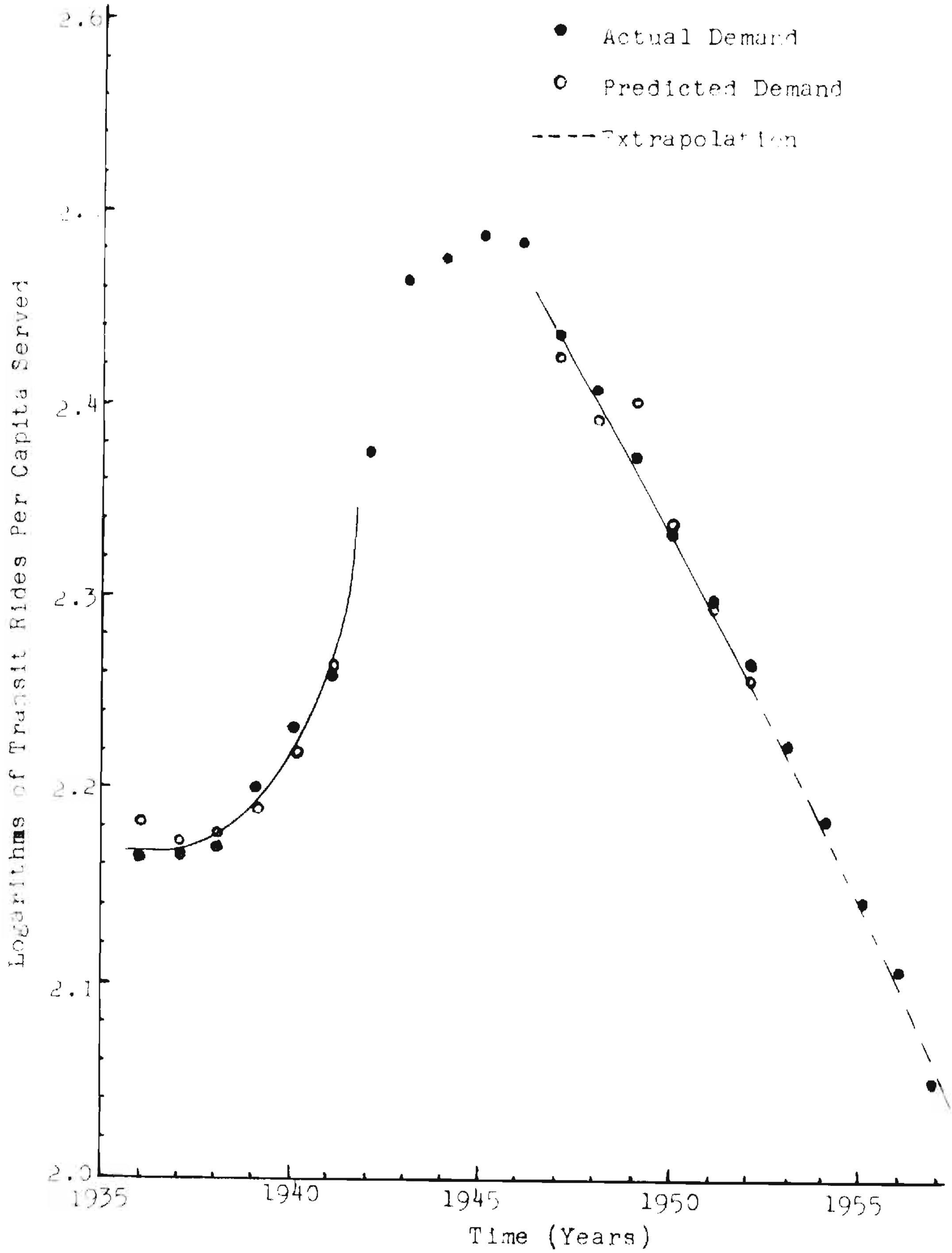


Fig. 7 : Extrapolation of Predicted Demand - Model III

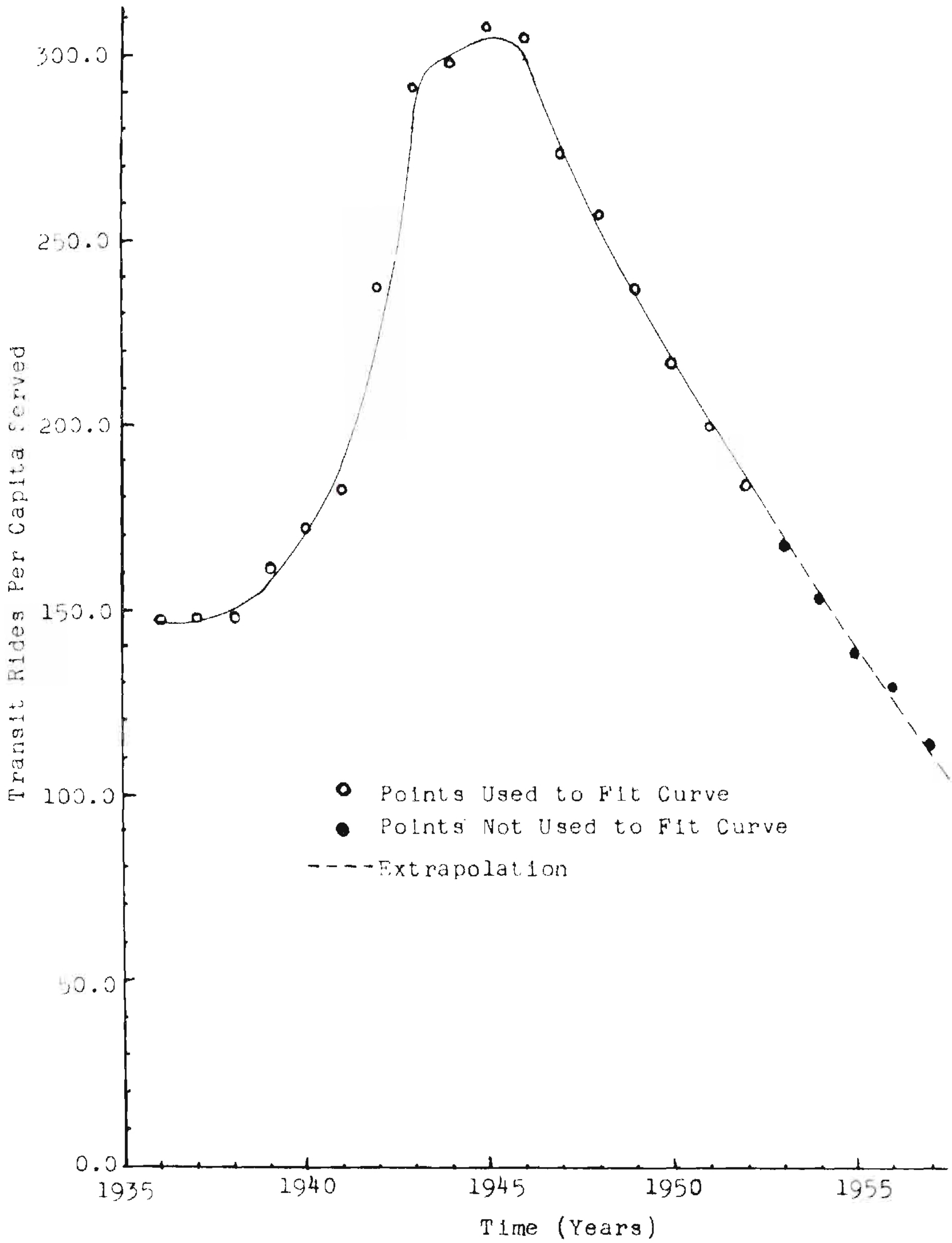


Fig. 8 : Relationship Between Actual Demand and Time

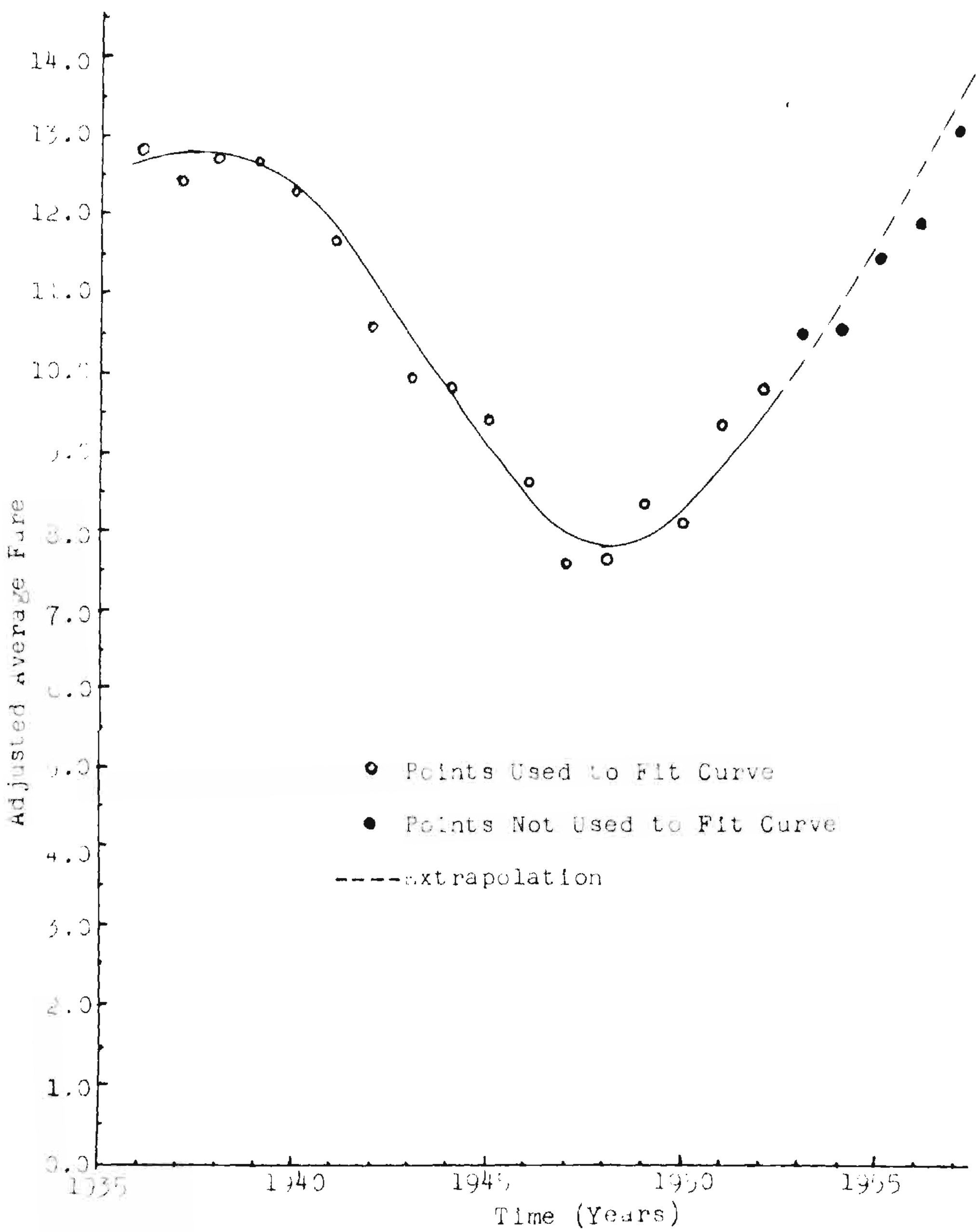


Fig. 9: Relationship Between Adjusted Average Fare and Time

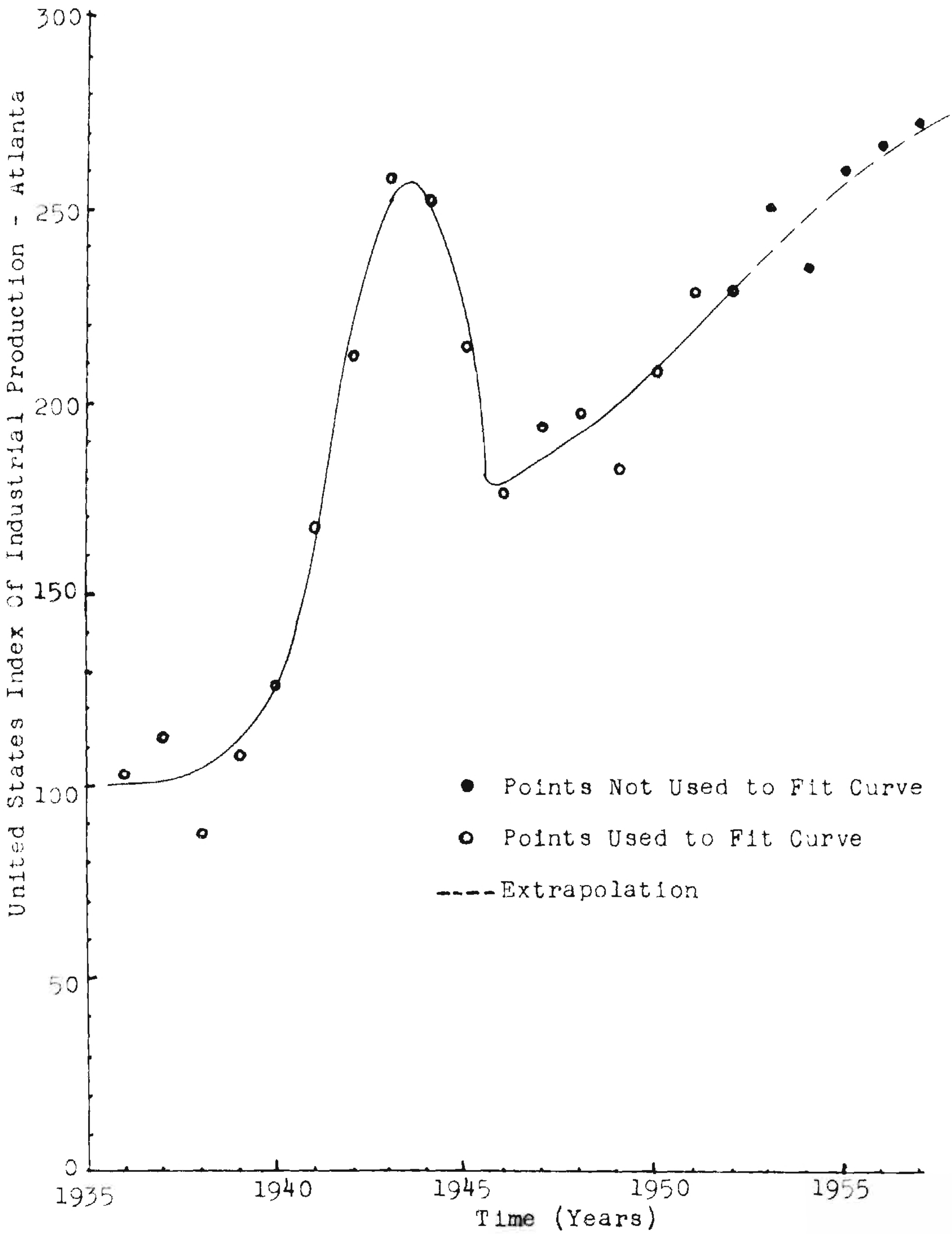


Fig. 10: Relation Between U.S. Index Of Industrial Production and Time

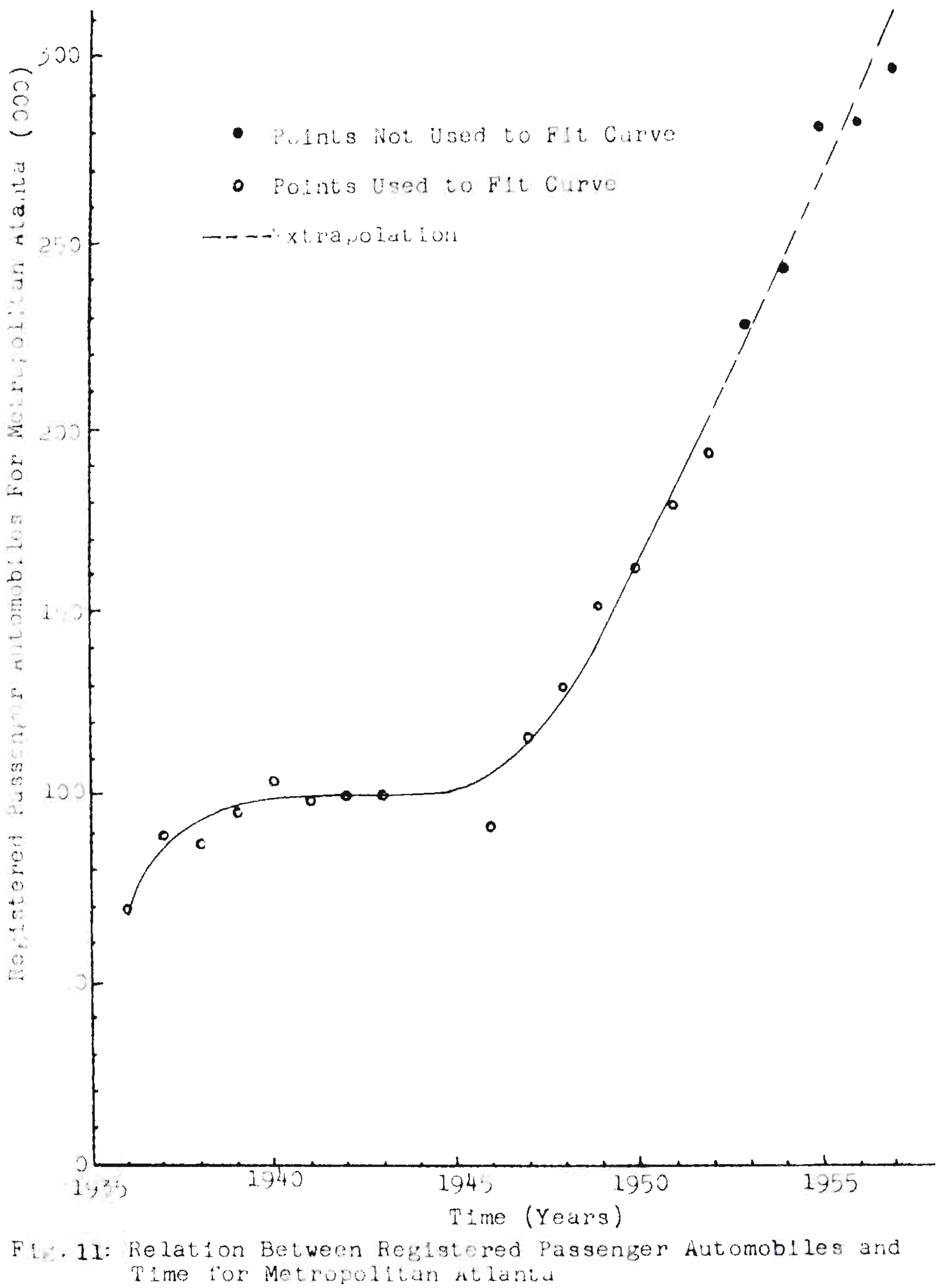


Fig. 11: Relation Between Registered Passenger Automobiles and Time for Metropolitan Atlanta

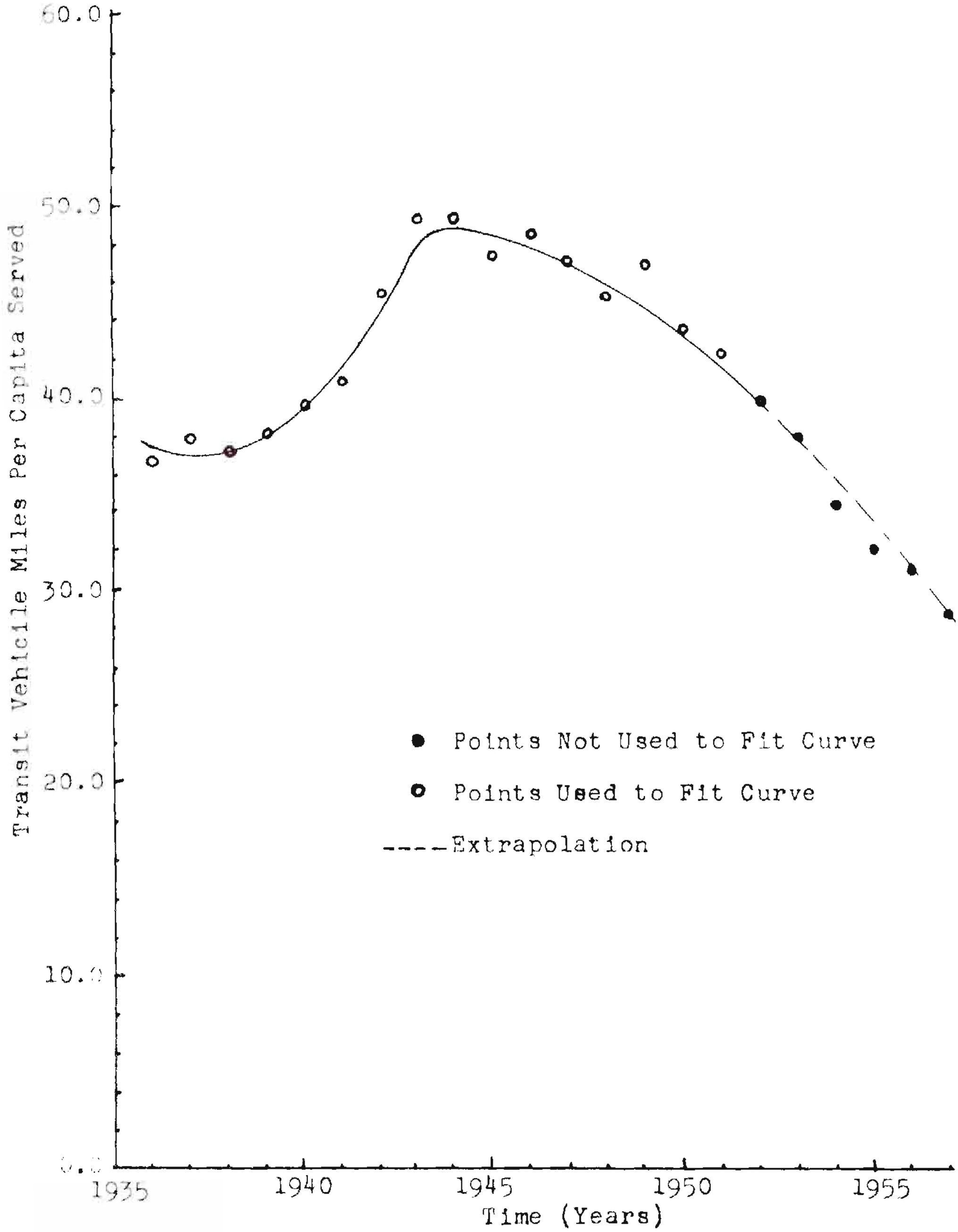


Fig. 12: Relation Between Transit Vehicle Miles Per Capita Served and Time

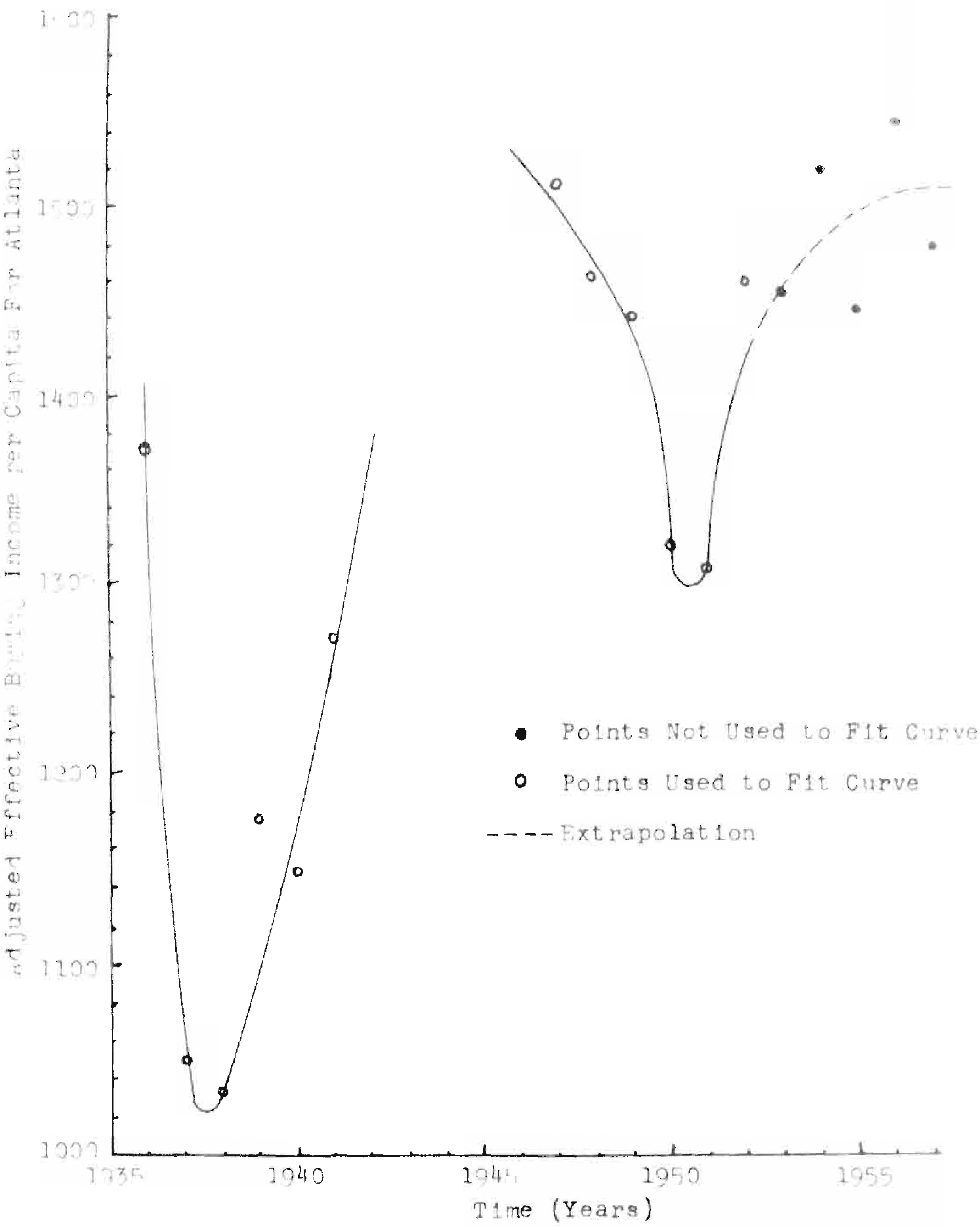


Fig.13: Relation Between Adjusted Effective Buying Income Per Capita For Atlanta and Time

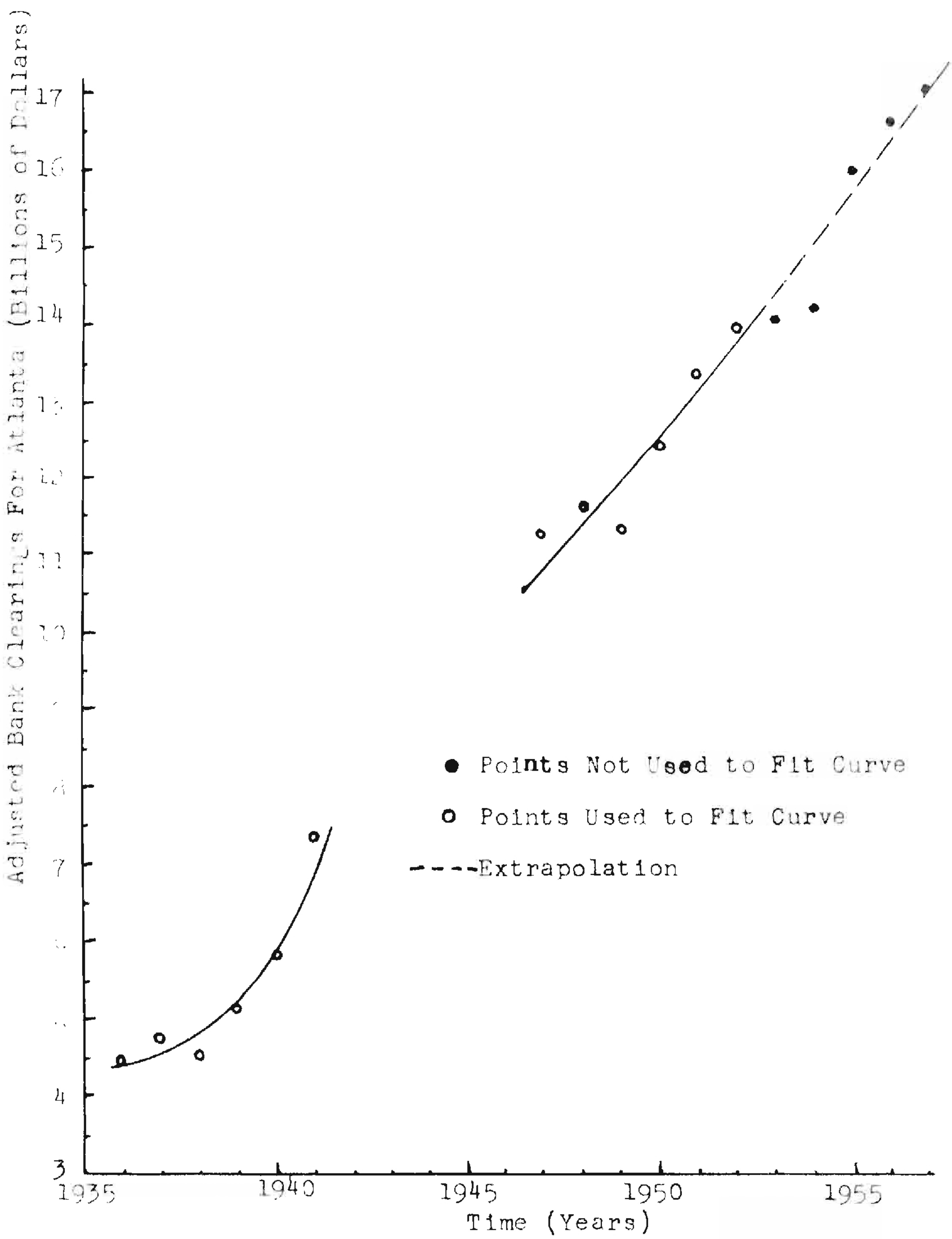
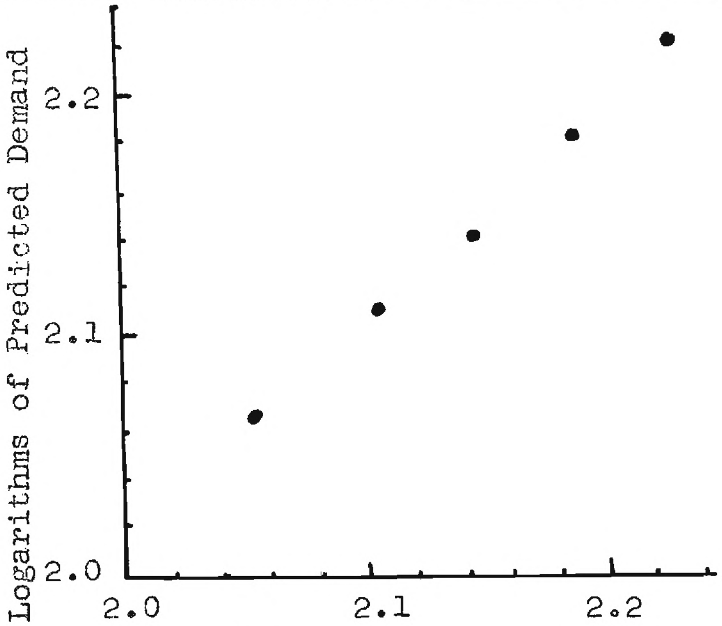


Fig. 14: Relation Between Adjusted Bank Clearings and Time

Table 20. Regression Estimates of Logarithms of Demand

Year	Model II	Model III	Model IV	Model V
1936	2.17058	2.18218	2.16921	2.19119
37	2.16007	2.17219	2.16012	2.17030
38	2.17767	2.17886	2.18385	2.18649
39	2.19788	2.19033	2.19444	2.18959
40	2.22614	2.22046	2.22398	2.21287
41	2.27761	2.26146	2.27944	2.25357
42	-	-	-	-
43	-	-	-	-
44	-	-	-	-
45	-	-	-	-
46	-	-	-	-
47	2.42608	2.42699	2.42522	2.41812
48	2.40003	2.39054	2.39928	2.38869
49	2.38542	2.40442	2.37965	2.39356
50	2.34441	2.33858	2.33964	2.34134
51	2.29188	2.29570	2.29307	2.29993
52	2.26363	2.25967	2.27329	2.27736
53	-	-	2.23008	2.23951
54	-	-	2.18341	2.18645
55	-	-	2.13101	2.12669
56	-	-	2.11939	2.11050
57	-	-	2.04696	2.04582

Predictive Accuracy by Extrapolation of Predicted Demand
Period 1953-1957



Model II (36-52, 9 Variables)

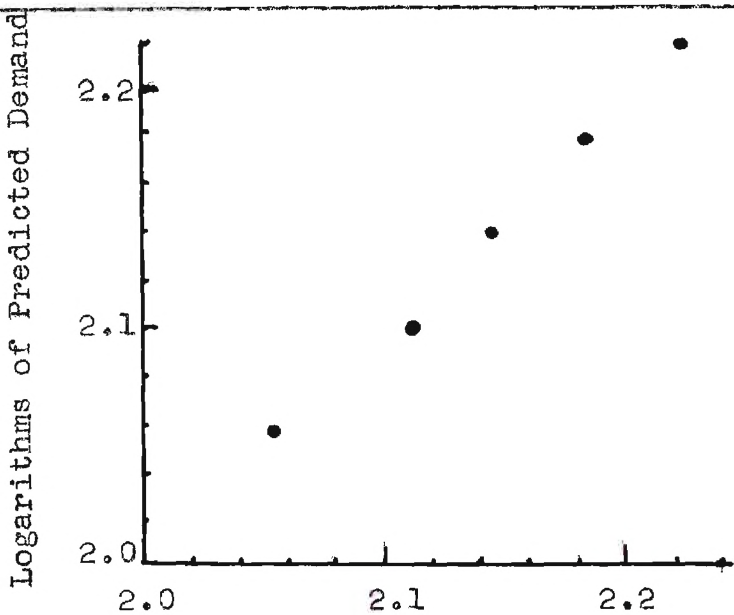
Year	Logarithms of Predicted Demand	Logarithms of Actual Demand
1953	2.225	2.2227
1954	2.185	2.1824
1955	2.145	2.1421
1956	2.106	2.1103
1957	2.067	2.0531

$$r = \frac{n\sum xy - \sum x \sum y}{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}$$

$$r = \frac{114.98452750 - 114.90331680}{(115.16800000-115.08998400)(114.80200000-114.71700000)}$$

$$r = \frac{0.081210700}{0.081559257} = \underline{\underline{0.995726}}$$

Predictive Accuracy by Extrapolation of Predicted Demand
Period 1953-1957



Logarithms of Actual Demand

Model III (36-52, 5 Variables)

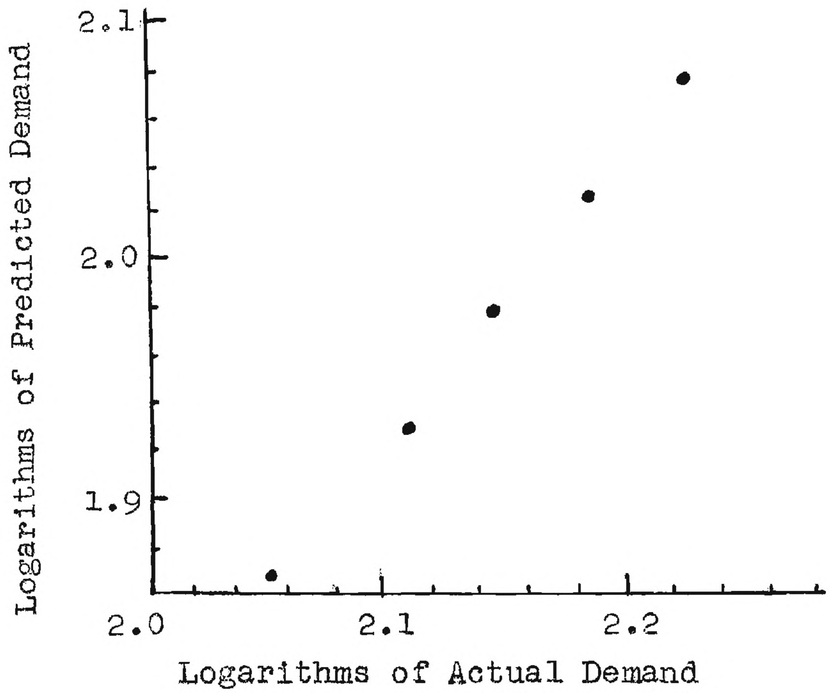
Year	Logarithms of Predicted Demand	Logarithms of Actual Demand
1953	2.218	2.2227
1954	2.178	2.1824
1955	2.139	2.1421
1956	2.099	2.1103
1957	2.056	2.0531

$$r = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}$$

$$r = \frac{114.57930500 - 114.496314000}{(114.357330-114.27610000)(114.80200000-114.71700000)}$$

$$r = \frac{0.082991000}{0.083768616} = \underline{\underline{0.990717}}$$

Predictive Accuracy by Extrapolation of Each Variable and Calculating Predicted Demand with the use of the Models
Period 1953-1957



Model II (36-52, 9 Variables)

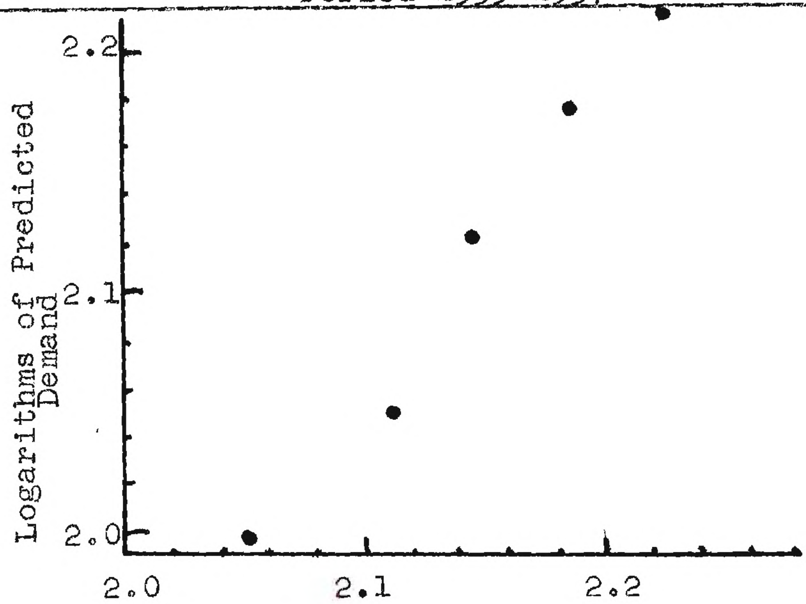
Year	Logarithms of Predicted Demand	Logarithms of Actual Demand
1953	2.076126	2.2227
1954	2.030594	2.1824
1955	1.980808	2.1421
1956	1.924833	2.1103
1957	1.865386	2.0531

$$r = \frac{n\sum xy - \sum x \sum y}{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}$$

$$r = \frac{105.90530705 - 105.79659702}{(97.70929250-97.56988580)(114.80221580)-(114.71695236)}$$

$$r = \frac{0.10871030}{0.10902400} = \underline{\underline{0.997123}}$$

Predictive Accuracy by Extrapolation of Each Variable and Calculating Predicted Demand with the use of the Models Period 1953-1957



Model III (36-52, 5 Variables)

Year	Logarithms of Predicted Demand	Logarithms of Actual Demand
1953	2.226442	2.2227
1954	2.175603	2.1824
1955	2.120746	2.1421
1956	2.057176	2.1103
1957	1.991213	2.0531

$$r = \frac{n\sum xy - \sum x \sum y}{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}$$

$$r = \frac{113.34508275 - 113.22368051}{(111.92379000)-(111.74984659)(114.80221580)-(114.71695236)}$$

$$r = \frac{0.12140224}{0.12178265} = \underline{\underline{0.996876}}$$

BIBLIOGRAPHY

Literature Cited

1. Baptiste, Jean, Catèchisme d'economic politique, 4th Edition, Paris, 1835, chapter xi, p. 104.

2. Cournot, Augustin, Researches into the Mathematical Principles of the Theory of Wealth, translation, Bacon, New York, 1897, p. 47.

3. Marshall, Alfred, Principles of Economics, 8th Edition, London, 1920, Book III.

4. Walras, Leon, Elements d'economie politique pure, 4th Edition; Lausanne and Paris, 1900, pp. v-vii and especially M. 1, p. vii; also, pp. 122-33 and 208-15.

5. Edgeworth, F. Y., Mathematical Psychics, London, 1881, p. 31.

6. Schultz, Henry, The Theory and Measurement of Demand, Chicago, 1938, p. 50.

7. Moore, H. L., Economic Cycles: Their Law and Cause, New York, 1914.

8. Schultz, op. cit., p. 65.

9. Smith, B. B., "The Error in Eliminating Secular Trends and Seasonal Variation Before Correlating Time Series", Journal of American Statistical Association, XX (1925), pp. 543-545.

10. Leontief, Wassily, "Ein Versuch zur statistichen Analyse von Angebot und Nachfraze," Weltwirtschaftliches Archiv, XX, Heft 1 (July, 1929), pp. 1-53.

11. Schultz, op. cit., pp. 93-95.

12. Pigou, A. C., "A Method of Determining the Numerical Value of Elasticities of Demand", Economic Journal, XX (1910) pp. 636-40; and reprinted in his Economics of Welfare (London 1920).

13. Pigou, A. C., "The Statistical Derivation of Demand Curves", op. cit., pp. 384-400, and reprinted in Pigou and Robertson, Economic Essays and Addresses, London, 1931, pp. 62-83.

14. *Ibid.*, p. 384.

15. Schultz, *op. cit.*, pp. 102-104.

16. Schultz, *op. cit.*, p. 110.

17. Frisch, Ragnor, *New Methods of Measuring Marginal Utility*, ("Beitrage zur okonomischev Theorie," No. 3, Tubingen, 1932).

18. Marschak, Jacob, *Elastizitat der Nachfrage* ("Beitrage zur okonomischen Theorie", No. 2, Tubungen, 1931).

19. Schultz, *op. cit.*, pp. 117-18.

20. Roy, René, "La Demande dans ses rapports avec la répartition des revenues", *Metron*, VIII, No. 3 (1930), pp. 101-153; "Les Lois de la demande", *Revue d'ēconomie politique*, XLV (1931), pp. 1190-1218.

21. Pareto, Vilfredo, "La Legga della domando", *Giornale degli economisti*, X (2nd series, 1895), pp. 59-68.

22. *A Plan For Transit Improvement in the Metropolitan Area of Atlanta*, Georgia State Senate Transit Study Committee, and Simpson and Curtin, Transportation Engineers and Consultants, November 1953.

23. Schultz, *op. cit.*, pp. 184-186.

24. Schultz, op. cit., Appendix C.

25. Hotelling, Harold, "Some New Methods in Matrix Calculation," *The Annals of Mathematical Statistics*, Vol. XIV, 1943, pp. 14-16.

26. International Business Machines' Multiple Regression Routine ST01. Phases I-IV, Rich Electronic Computer Center.

27. Schultz, *op. cit.*, p. 190.

28. Moore, H. L., *Synthetic Economics* (New York, 1929), p. 55.

29. Schultz, *op. cit.*, p. 214.

30. Schultz, *op. cit.*, p. 198.

31. Bennett and Franklin, *Statistical Analysis in Chemistry and the Chemical Industry*, New York, 1954, p. 287.